TESIS DE DOCTORADO

Variability and function of immune genes in the mussel *Mytilus galloprovincialis*

Magalí Rey Campos

ESCUELA DE DOCTORADO INTERNACIONAL DE LA UNIVERSIDAD DE SANTIAGO DE COMPOSTELA (EDIUS)

PROGRAMA DE DOCTORADO EN CIENCIAS MARINAS, TECNOLOGÍA Y GESTIÓN

SANTIAGO DE COMPOSTELA

2021

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DEL AUTOR DE LA TESIS

Variability and function of immune genes in the mussel *Mytilus galloprovincialis*

Dña. Magalí Rey Campos

Presento mi Tesis Doctoral, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

1) *La Tesis abarca los resultados de la elaboración de mi trabajo.*
2) *En su caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
3) *La Tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.*
4) *Confirma que la Tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por ella para la obtención de otros títulos.*

En , a de de 2021

Fdo: Magalí Rey Campos

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DE LOS DIRECTORES DE LA TESIS

Variability and function of immune genes in the mussel *Mytilus galloprovincialis*

D. Antonio Figueras Huerta y Dª. Beatriz Novoa García

INFORMA/N:

Que la presente tesis, se corresponde con el trabajo realizado por Dª. Magalí Rey Campos, bajo nuestra dirección, y autorizamos su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como directores de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.

De acuerdo con lo indicado en el Reglamento de Estudios de Doctorado, declaramos también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de Monográfica con reproducción de publicaciones, en los que la participación del doctorando/a fue decisiva para su elaboración y las publicaciones se ajustan al Plan de Investigación.

En , a de de 2021

Fdo: Antonio Figueras Huerta

Fdo: Beatriz Novoa García

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

With the support of the Spanish Ministerio de Ciencia e Innovación through a FPI research Grant (BES-2016-076302) associated to the project (AGL2015-65705-R). This work has also been supported by the VIVALDI project (678589; EU H2020).

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Agradecimientos

En este “capítulo” de la tesis, me gustaría dar un GRACIAS enorme a todos los que me han ayudado en esta etapa.

Primero me gustaría dar las gracias a mis directores, Bea y Antonio. Me habéis dado la oportunidad de trabajar en vuestro grupo y de seguir formándome. He aprendido muchísimo durante estos años, y estoy segura de que me ayudará en mi futuro profesional. Recuerdo que la primera vez que entré en el IIM fue para la entrevista de la FPI. Por aquel entonces ya sabía que esta etapa sería exigente, y así ha sido, pero ahora sé que con constancia y perseverancia se alcanzan los objetivos. Esto también lo he aprendido aquí. Espero haberos devuelto la confianza y haber estado a la altura. En definitiva, para vosotros el primer GRACIAS de este “capítulo”.

La segunda parada de estos agradecimientos es para el grupo, “Inmunología y Genómica”. Se me ocurren numerosas ocasiones en las que he necesitado de alguno de vosotros. Detrás de cada experimento, resultado o publicación está la experiencia o ayuda de mis compañeros. A Rebe tengo mucho que agradecerle…empezar nunca es sencillo y tú me lo pusiste en bandeja…me has ayudado en todo lo que he necesitado… ¡Muchísimas gracias por todo! Con Álex he aprendido a usar el citómetro, a hacer inmunos…y a todo lo que le he pedido, que no ha sido poco… ¡Gracias Álex! Bego y Raquel Aranguren siempre han estado dispuestas a ayudar con la histología…por eso y por el afecto, ¡Gracias! Con Marimar he trabajado al principio de esta etapa y también he aprendido mucho… ¡Muchas gracias MM! Patri me ha resuelto muchas dudas…porque diez años en el laboratorio, son muchos años… ¡Gracias por todo Patri! Marta Sendra es nuestra gaditana experta en nanoplásticos… y en hacernos reír… ¡Muchas gracias Marta! Y si surgen dudas sobre el metabolismo, ahí está Marta Librán…que además siempre tiene una anécdota graciosa que contar… ¡Gracias por todo Marta! Judit, siempre está dispuesta a ayudar…zebras, células, inmunos, confocal…sabes mucho Judit… ¡Mil gracias por todo! Y mención especial para mis compañeros “predocs”. Marga, Raquel Lama, Raquel Ríos, Antonella y Amaro…entre todos es más fácil aclararse con los trámites, la redacción

de la tesis, los seminarios, los cursos…pero lo más importante es el apoyo moral y el afecto...GRACIAS.

Aunque no sólo ha habido tiempo para los experimentos y los papers. Antes de que la pandemia nos lo quitase, compartíamos "cafés", comidas y reuniones de grupo (...sí, para que quede constancia, en 2020 hizo su aparición estelar el SARS-CoV-2...no sabemos aún cuándo terminará...pero por si algún día se nos llega a olvidar, está bien que quede constancia...). Hemos comentado, debatido y bromeado sobre multitud de temas de actualidad. Es probable que no se nos haya quedado ninguno en el tintero. Así que, en este sentido, objetivo también más que cumplido. ¡Mi más sincero agradecimiento para todo el grupo!

I would also like to thank Alberto Pallavicini and Marco Gerdol for allowing me to do my predoctoral stay in their group. I learned a lot and the stay was really productive. They and all the members of the group were very kind with me. I really felt at home during the 3 months I spent in Trieste. Thank you very much for everything! Grazie mille!

También me gustaría agradecer el apoyo y el afecto a mis amigos y familia. En especial a Fátima y Nerea, con vosotras compartí etapa universitaria… y todo lo que vino después ¡Muchas gracias! También a Vero, por la confianza y el respeto… ¡Muchas gracias por todo! A Fani, por el cariño y el apoyo… ¡Una vez más, un millón de gracias!

Y por último a mis padres y mi hermana, por haber sido un soporte incondicional ¡GRACIAS!

Resumo

Variabilidade e función de xenes inmunes do mexillón *Mytilus galloprovincialis*

O mexillón mediterráneo (*Mytilus galloprovincialis*) é un molusco bivalvo de gran interese comercial en todo o mundo, xa que se producen máis de 2 millóns de toneladas ao ano. Ademáis dos beneficios que aporta ao sector acuícola, é moi interesante dadas algunhas características da súa bioloxía. Estes animais viven fixados ás rocas, polo tanto, practicamente inmóbiles e teñen un modo de vida filtrante a través do que obteñen o alimento. Estas características fan que sexan especialmente vulnerables, xa que están constantemente expostos a todo tipo de patóxenos. Aínda así, non se rexistraron eventos de mortalidade masiva no medio natural (algo que sucedeu con outras especies de bivalvos como ostras ou ameixas). Esta resistencia parece residir na gran variabilidade de moléculas inmunes que codifica o seu xenoma, dende os receptores de patóxenos ata os péptidos antimicrobianos efectores. Nesta tese analizouse a resposta transcriptómica mediante *RNA-Seq* das principais células inmunes do mexillón, os hemocitos, despois dunha infección cunha bacteria patóxena, *Vibrio splendidus*. Ademáis, dada a crecente información sobre a existencia dunha memoria innata en invertebrados (estes animais carecen de inmunidade adaptativa e anticorpos), tamén se estudou a resposta dos hemocitos tras dúas exposicións sucesivas ao mesmo patóxeno, empregando a mesma tecnoloxía. Estes estudos proporcionaron pistas sobre a función de numerosos xenes. Entre eles destacan as miticinas, uns péptidos clave para a defensa destes animais e cuxa función clásica antibacteriana e antiviral se ampliou nesta tese, comprobando a súa actividade quimiotáctica e rexenerativa. Ademais, o estudo da secuencia do xene descubriu a enorme variabilidade xenómica das miticinas.

Palabras clave: mexillón, *RNA-Seq*, resposta inmune, miticinas, variabilidade.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Resumen

Variabilidad y función de genes inmunes del mejillón *Mytilus galloprovincialis*

El mejillón mediterráneo (*Mytilus galloprovincialis*) es un molusco bivalvo de gran interés comercial en todo el mundo, ya que se producen sobre 2 millones de toneladas al año. Además de los beneficios que aporta al sector de la acuicultura, tiene mucho interés dadas algunas características de su biología. Estos animales viven anclados a las rocas, prácticamente inmóviles, lo que condiciona su modo de vida filtrador. Estas características los hacen especialmente vulnerables, ya que están expuestos de manera constante a todo tipo de patógenos. Aun así, no se han registrado eventos de mortalidad masiva en el medio natural (algo que sí ha ocurrido con otras especies de bivalvos como la ostra o la almeja). Esta resistencia inmunológica parece residir en la gran variabilidad de moléculas inmunes que codifica su genoma, desde receptores de patógenos hasta péptidos efectores antimicrobianos. En esta tesis se ha analizado la respuesta transcriptómica por medio de *RNA-Seq* de las principales células inmunes del mejillón, los hemocitos, tras una infección con una bacteria patógena, *Vibrio splendidus*. Además, dada la creciente información acerca de la existencia de una memoria innata en invertebrados (estos animales carecen de inmunidad adaptativa y anticuerpos), también se estudió, la respuesta de los hemocitos tras dos exposiciones sucesivas al mismo patógeno. Estos estudios aportaron pistas sobre la función de numerosos genes. Entre ellos se encuentran las miticinas, unos genes clave para la defensa de estos animales y cuya función clásica antibacteriana y antiviral se ha ampliado en esta tesis, habiéndose probado su actividad quimiotáctica y regeneradora. Además, el estudio de la secuencia del gen ha destapado la enorme variabilidad genómica de las miticinas.

Palabras clave: mejillón, *RNA-Seq*, respuesta inmune, miticinas, variabilidad.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Abstract

Variability and function of immune genes in the mussel *Mytilus galloprovincialis*

The Mediterranean mussel (*Mytilus galloprovincialis*) is a bivalve mollusc of great commercial interest worldwide, since it is produced over 2 million tons per year. But, in addition to the benefits it brings to the aquaculture sector, it is of great interest given some of the characteristics of its biology. These animals live anchored to the rocks, therefore, practically immobile and have a filtering way of life through which they get food. These characteristics make them especially vulnerable, since they are constantly exposed to all kinds of pathogens. Even so, no mass mortality events have been recorded in the natural environment (something that has occurred with other bivalve species such as oysters or clams). This immune resistance seems to reside in the great variability of immune molecules that their genome encodes, from pathogen receptors to antimicrobial effector peptides. The transcriptional response of hemocytes, the main immune cells of mussels, have been studied by RNA-Seq, in an infection context with a pathogenic bacterium such as *Vibrio splendidus*. Moreover, given the increasing information about the existence of an innate memory in invertebrates (these animals lack adaptive immunity and antibodies), the response of hemocytes after two successive exposures to the same pathogen was also studied using the same technology. These studies provided clues about the function of numerous genes. For instance, the myticins, key genes for the defense of these animals and whose classic antibacterial and antiviral function has been extended in this thesis, having been proved its chemotactic and regenerative activity. In addition, the study of the gene sequence has revealed the enormous genomic variability of myticins.

Keywords: mussel, RNA-Seq, immune response, myticins, variability.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

List of scientific publications and quality criteria of the journals

The results exposed in this thesis have been published in the following scientific publications:

Chapter 2: Rey-Campos, M.; Moreira, R.; Valenzuela-Muñoz, V.; Gallardo-Escárate, C.; Novoa, B.; Figueras, A. High individual variability in the transcriptomic response of Mediterranean mussels to *Vibrio* reveals the involvement of myticins in tissue injury. Sci Rep. 9, 3569 (2019). doi:10.1038/s41598-019-39870-3. (Open Access).

Author contribution: mussel infections, sampling and RNA extraction, analysis of the generated data and drafting of the manuscript.

Chapter 3: Rey-Campos, M.; Moreira, R.; Gerdol, M.; Pallavicini, A.; Novoa, B.; Figueras, A. Immune tolerance in *Mytilus galloprovincialis* hemocytes after repeated contact with *Vibrio splendidus*. Front Immunol.10, 1894 (2019). doi:10.3389/fimmu.2019.01894. (Open Access).

Author contribution: mussel infections, sampling and RNA extraction, analysis of the generated data, functional analysis and drafting of the manuscript.

Chapter 4: Rey-Campos, M.; Moreira, R.; Romero, A.; Medina-Gali, R.M.; Novoa, B.; Gasset, M.; Figueras, A. Transcriptomic analysis reveals the wound healing activity of mussel myticin C. Biomolecules. 10, 133 (2020). doi:10.3390/biom10010133. (Open Access).

Author contribution: experimental procedures, sampling and RNA extraction, analysis of the generated data and drafting of the manuscript.

Chapter 5: Rey-Campos, M.; Novoa, B.; Pallavicini, A.; Gerdol, M.; Figueras, A. Comparative genomics reveals a significant sequence variability of myticin genes in *Mytilus galloprovincialis*. Biomolecules. 10, 943 (2020). doi:10.3390/biom10060943. (Open Access).

Author contribution: methodology and formal analysis and drafting of the manuscript.

Journal	IF 2019	5-year IF	Quartile	Rank	ISSN
Scientific Reports	3.998	4.324	Multidisciplinary Science (Q1)	17/71	2045-2322
Frontiers in Immunology	5.085	5.487	Immunology (Q1)	38/158	1664-3224
Biomolecules	4.082	4.388	Biochemistry & Molecular Biology (Q2)	98/297	2218-273X

Other scientific publications:

Taboada, X.; **Rey, M.**; Bouza, C.; Viñas, A. Cytogenomic analysis of several repetitive DNA elements in turbot (*Scophthalmus maximus*). Gene. 644, 4–12 (2018). doi: 10.1016/j.gene.2017.12.013.

Sendra, M.; Saco, A.; **Rey-Campos, M.**; Novoa, B.; Figueras, A. Immune-responsive gene 1 (IRG1) and dimethyl itaconate are involved in the mussel immune response. Fish Shellfish Immunol. 106, 645–655 (2020). doi:10.1016/j.fsi.2020.07.034.

Saco, A.; **Rey-Campos, M.**; Novoa, B.; Figueras, A. Transcriptomic response of mussel gills against *Vibrio splendidus* reveals its role in the immune response. Front Immunol. 11, 3273 (2020). doi: 10.3389/fimmu.2020.615580.

Moreira, R.; Romero, A.; **Rey-Campos, M**.; Pereiro, P.; Rosani, U.; Novoa, B.; Figueras, A. Stimulation of *Mytilus galloprovincialis* hemocytes with different immune challenges induces differential transcriptomic, miRNomic and functional responses. Front Immunol. 11, 3318 (2020). doi:10.3389/fimmu.2020.606102.

Rey-Campos, M.; Novoa, B.; Pallavicini, A.; Gerdol, M.; Figueras, A. Comparative Genomics Reveals 13 Different Isoforms of Mytimycins (A-M) in *Mytilus galloprovincialis*. Int J Mol Sci. 22, 3235 (2021). doi:10.3390/ijms22063235.

Index

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 1

GENERAL INTRODUCTION AND OBJECTIVES

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 1: General introduction and objectives

1.1 GENERAL INTRODUCTION

1.1.1 *Mytilus galloprovincialis* biology

Mytilus galloprovincialis is a bivalve mollusc characterized by a solid shell, equivalve and with a roughly triangular profile. These shells hinge to each other allowing opening for feeding and closing for protection from predators or drying out at low tide. From an anatomical perspective, gills are the most exposed tissue when opening the two shells. Gills constitute the respiratory system of the mussel and are formed by laminar structures which, in turn, are comprised of filaments joined together. Just below the gills it can be found the mantle, which covers the whole mussel body. The mantle encloses gametes, white for spermatozoids and orange for oocytes, that are released during spawning events for the development of fertilized eggs that will become free-swimming larvae capable of reaching great distances. After several larval stages, a metamorphosis process occurs and the juveniles, now similar to adults, attach to a substrate to continue the life cycle. The organ used to dig and hold onto the substrate is a foot, which is reddish and muscular and is located below the visceral mass. This locomotive organ contains a gland towards the end, that secretes the byssus, formed by a set of filaments of protein nature used by the mussel to anchor itself to the rocks. Moreover, mussels contain some muscles in charge of closing shells, being the largest the posterior adductor. All these tissues are shown in the figure 1.

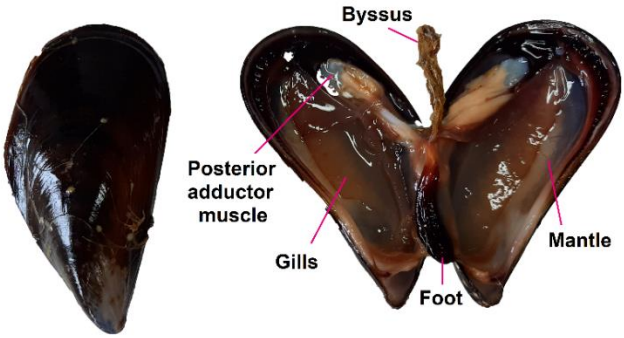

Figure 1. Image of the mussel *Mytilus galloprovincialis*, showing the main anatomical parts. Image source: own work.

1.1.2 *Mytilus galloprovincialis* as an aquaculture resource

In the last 50 years, there has been a dramatic growth in the human world's population (Figure 2). This fact is essentially due to the increase in life expectancy as well as urbanization processes and migratory movements. According to the United Nations, the world population is expected to increase by two billion people in the next 30 years, reaching 10 billion people in 2050 [1].

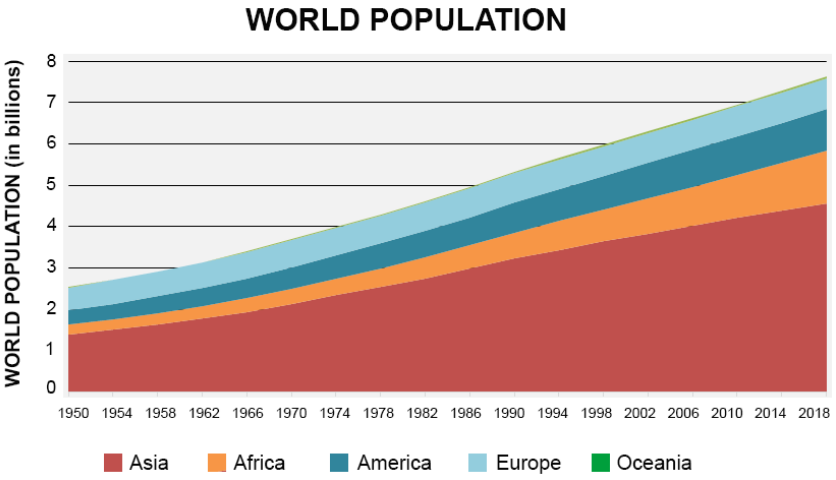

Figure 2. World population data from 1950 to 2018. Data source: United Nations. Graph: own work.

This fact implies a growing increase in food needs, so fisheries and aquaculture have become very important in recent years. Fish, crustaceans, molluscs and other aquatic animal's capture fisheries and aquaculture reached about 180 million tons in 2018, 87% of which were used for human consumption (Figure 3). China leads the productive values with 35% of the global fisheries and aquaculture production [2].

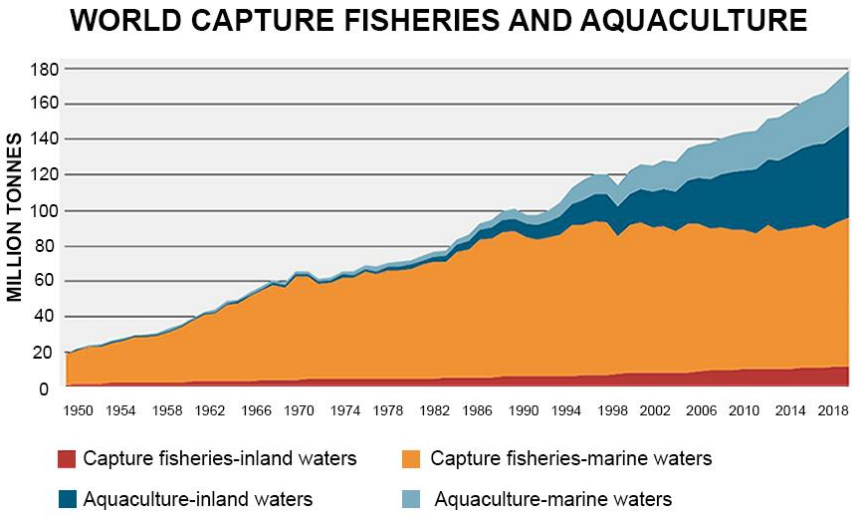

Figure 3. World capture fisheries and aquaculture data from 1950 to 2018. Graph source: FAO (Open Access).

In terms of human consumption, food fish has increased at an average annual rate of 3%, higher than other animal protein foods as meat or milk, that increased by 2% per year.

In 2018, the global aquaculture production reached 82 million tons, of which 17.7 million were molluscs, mainly bivalves, and representing the 56.3% of the production of marine aquaculture. The most produced bivalve molluscs are mussels, clams, scallops and oysters [2].

Mussels are present worldwide and the global production of this bivalve reached 2 million tons in 2017 [2]. Although the production numbers per country are approximate due to the fact that reports do not

collect clearly differentiated production data for each mussel species. China would produce about half, while Chile and Spain would produce around 230,000 tons each per year. Galicia is responsible for almost all the Spanish production. The culture of this mollusc is developed on rafts, floating platforms that have become part of the landscape of the Galician coast (Figure 4).

Figure 4. Image of the rafts in the Ría de Vigo, Galicia. Image source: own work.

1.1.3 Mussel immune system

In addition to its economic value, *M. galloprovincialis* has shown a great biological value in recent years. From an ecological point of view, its ubiquity has caused it to be considered an invasive species [3] and, the sessile condition as well as the filter-feeding way of life explain why these animals serve as ecological markers of pollution [4,5]. Precisely, existing naturally anchored to the rocks, that is, without being able to move or escape, and living in a marine environment, persistently filtering water containing pathogens and contaminants, could be contradictory to their widespread geographical distribution. However, no information about massive mortalities in naturally conditions have been reported [6,7], unlike other species of the same class, such as clams and oysters [8,9]. The explanation of this great capacity to resist adverse conditions in the environment would be explained by the features of the immune system of mussels. As invertebrates, these animals lack adaptive immunology, therefore, they have to combat the infections based on their innate immunity. Hemocytes are the central immune cells of mussels, that recognize conserved pathogen-associated molecular patterns (PAMPs), via pattern recognition receptors (PRRs) [10,11], as well as danger signals, and danger/damage-associated molecular patterns (DAMPs) [11,12]. After recognition processes, these cells act by chemotaxis, encapsulation, phagocytic activity and releasing oxygen and nitrogen radicals, as well as activating intracellular signaling pathways to finally trigger the synthesis of antimicrobial effectors [13,14].

1.1.3.1 Pattern recognition receptors (PRRs)

Several PRRs have been defined in mussels [15]. These receptors are extracellular, membrane-bound or cytosolic molecules and they are characterized by a range of domains, usually carbohydrate binding proteins. For instance, **C1q domain containing proteins** constitute a large class of extracellular PRRs first discovered in *M. galloprovincialis* in 2009 by Venier et al. [16]. A significant expansion has been reported in some bivalves such as *M. galloprovincialis* or *Crassostrea gigas* [17,18], being possible to explain the broad spectrum

of pathogens recognized by the C1q domain (bacteria, fungi, parasites…). Expression of C1q domain containing protein is widespread in most mussel tissue as well as during different ontogeny stages and in naïve and infected conditions, so the exact function of these proteins, apart from recognition, remains undetermined.

Other kind of extracellular pathogen receptors are the molecules containing a **C-type lectin domain**, that are able to trigger the lectin pathway of the complement system. As C1q domain containing proteins, these C-type lectins show a significant abundance in *M. galloprovincialis* [15]. As in mussels, in some other bivalves such as *Ruditapes philipinarum* and *Crassostrea gigas*, C-type lectin receptors would participate in the pathogen identification process and the immune defense response [19,20].

Fibrinogen-related proteins (FREPs) have been identified in invertebrates as players in the defense response against pathogens (unlike the coagulation function they have in vertebrates [21]). Again, a variable repertoire of this secreted molecule was identified in *M. galloprovincialis*, showing a significant increase of the expression after bacterial infection or PAMPs treatment and displaying opsonic activities [22].

Apart from soluble receptors, immune cells also display PRRs bounded in their membranes. It is the case of **Toll-like receptors (TLRs**) that have a conserved intracellular Toll-interleukin-1-receptor (TIR) domain, a transmembrane region and an extracellular region of leucine-rich repeats (LRRs), capable of binding bacterial components as well as viral RNA [15]. After recognition, TLRs initiate the activation of transcription factors as NF-kB and interferon-regulatory factors (IRFs) that will trigger the expression of pro-inflammatory cytokines and antiviral molecules. These receptors were discovered in *M. galloprovincialis* by Toubiana et al. in 2013 [23] but it is still not clear what specific stimuli they recognize or how they are distributed in hemocytes or other mussel cells.

In addition to the previously cited extracellular and membrane-bound receptors, some cytosolic receptors have also been identified in mussels: **NOD-like receptors (NLRs)** and **RIG-like receptors**

(RLRs) [15]. These receptors amplify the sensing of bacterial components and viral ds-RNA respectively, but more studies are necessary in *M. galloprovincialis*. Gerdol and Venier [15] reported a NLR-like sequence and a RIG-like receptor with canonical domain organization.

After recognition processes, a complex signaling cascade occurs, leading to a whole series of immune cellular processes and secretion of effector molecules.

1.1.3.2 Immune cellular processes

After a tissue damage and/or infection, immune cells are recruited to the damaged area to start the whole process of fighting against the pathogen or repairing the injured tissue. This process of cell migration is known as **chemotaxis** [24]. In mussels, as previously mentioned, hemocytes are the immune cells that travel looking for solve a compromised immune status [25]. Some new information about this hemocytes function is developed in further chapters of this thesis. After the cell migration, hemocytes are capable of encapsulating foreign particles and phagocyting them. **Phagocytosis** is a universal cell function, which starts recognizing and binding foreign particles and leads to its internalization and degradation [26]. This cellular process of pathogen engulfment is linked to the production of reactive oxygen species (**ROS**) and nitric oxide (**NO**), that are key in microbial killing [27]. The capability of mussel hemocytes to phagocyte particles has been widely demonstrated, not only recognizing and internalizing pathogens [28], but also nanoplastics [29].

1.1.3.3 Antimicrobial effectors

One of the keys of mussel resistance to disease is the great variety of immune effectors that they express in the genome. One of these effectors are the antimicrobial peptides (AMPs) that are a widespread group of molecules characterized by possess a structure of amino acids rich in cysteines and with antimicrobial properties. Among

the best known in mussels are **defensins** [30], **mytilins** [31], **mytimycins** [32] and **myticins** [33]. These molecules are composed by a signal peptide, followed by the mature Cys-rich regions and a C-terminal extension.

In terms of sequence variability, although mussel AMPs show quite inherent genomic variability, the case of the myticins is remarkable. Since the identification of forms A and B [33], a third form named as myticin C [34] and an extreme interindividual variability [35] have appeared in the last years. One of the chapters of this thesis focuses on the study of this feature.

Attending to the functional activity of AMPs, it is worth noting the antifungal activity of mytimycins, that clearly differs from mussel defensins, mytilins and myticins [32,36]. It is also interesting to mention how new functions, not strictly antimicrobial, have been identified in relation to myticin C. Specifically, a chemotactic activity defined in 2011 by Balseiro et al. [25] attributed to the myticin C a new chemokine role, which was confirmed in later works that are part of the present thesis, and that involved them in tissue injury and regeneration processes.

Other kind of molecular effectors are proteins as **lysozymes**, that hydrolyze 1,4-betaglycosidic bonds between N-acetylmuramic acid and N-acetylglucosamine in bacterial peptidoglycan. Some works in *M. galloprovincialis* have shown a strong modulation after PAMP treatments as well as after the injection of several mussel pathogenic bacteria [14,37]. Moreover, some **protease inhibitors,** key to counteract invading pathogen proteases, have been transcriptomically found, such as several tissue inhibitors of metalloproteinases (TIMPs) and Kazal-type protease inhibitors in mussels, but it is necessary to perform more functional studies.

1.1.3.4 Immune modulators

Cytokines are a large and heterogeneous group of regulatory molecules which comprise interleukins, interferons, tumor necrosis factor (TNF) and chemokines. In detail, only one pro-inflammatory

interleukin has been defined in *M. galloprovincialis* so far, **IL-17** [38]. Six different forms have been defined to date, and the exact function of each one is not yet known. However, the expression of several isoforms is specially up-regulated in mussel gills after an infection with a mussel pathogenic bacteria by bath (under experimental conditions similar to natural), indicating their relevance as immune modulators [39]. As far as **interferons** are concerned, there is no evidence that they exist in the mussel genome, however, almost all the components and molecules that belong to the IFN pathway have been found. Even interferon-induced proteins as IFI44L showed up-regulation associated to viral infections in bivalves such as *Crassostrea gigas* [40] as well as changes of expression in mussel hemocytes in the context of bacterial infections as it is reflected in later chapters of this thesis. Regarding **TNF**, some mussel transcripts were reported by Gerdol and Venier in 2015 [15] but there is little information about their function in mussels, however, they would have direct action on processes such as inflammation or apoptosis. Finally, as stated above, Balseiro et al., [25] proposed the myticin C as the first **chemokine-like protein** found in *M. galloprovincialis*. Nowadays, there is still very little information on this type of molecules capable of attracting in bivalves and more studies are needed.

1.1.4 Priming immunity in invertebrates

Because of the lack of an adaptive immune system and antibody production, it has been classically thought that invertebrates could not develop immune memory. But, in the last decade this idea has drastically change and a new form of innate immune memory has been defined as **immune priming** or **trained immunity** [41,42,43,44]. These concepts comprise the idea of an enhanced innate immune response after a previous stimulation with a non-self antigen. It has been found in many invertebrates, as well as their offspring, and it can occur against the same pathogen or against a different one. However, it is important to highlight that this heightened immunity is far from the specificity and amplification of adaptive vertebrate immunity. This

memory offers a new clue about how invertebrates can face pathogen in a successful manner [41].

In the last years, numerous examples of trained immunity have been defined and studied in invertebrates [45,46,47]. Also, in molluscs, as it is the case of *Crassostrea gigas*, in which an important effort has been made to understand the mechanism. Specifically, a poly I:C stimulation and a subsequent infection with OsHV-1 induced an antiviral state in oysters and it could be ascertained that the protection was viral-specific, since the priming using heat-killed *Vibrio splendidus* did not provide protection against OsHV-1 infection [48]. In a later work, this protection was confirmed and it was observed that could remain for up to 5 months [49,50]. Moreover, offspring produced from poly I:C-treated parents had double chances of surviving an exposure to OsHV-1 [51]. All information of these studies that relates a poly I:C stimulation and the trained response after the infection with OsHV-1 is reviewed in Green and Speck, 2018 [52].

This phenomenon has not been extensively studied in *M. galloprovincialis*, but this idea of a modified immune response after two encounters with the same pathogen has been considered and analyzed in a later chapter of this thesis.

1.1.5 Genome and transcriptome resources

The available genomic and transcriptomic resources have increased in recent years, being recently published the mussel genome [53]. The most interesting finding of this relevant work is the pan-genomic structure, with a core set of genes (represented by the 75% of genes), shared by all the individuals; and dispensable genes (about 25% of genes) that can be present or absent depending on the individual. This phenomenon is known as **presence/absence variation (PAV)** and may explain the great diversity as well as the marked adaptive capabilities of mussels [53]. Related to this, the resources provided by the genome sequencing together with the variability of mussel AMPs previously defined enabled the analysis of the variability of myticins, the last chapter of this thesis.

In addition to the sequencing of the genome, several transcriptomic studies analyzing different tissues as well as different subjects and experimental conditions have been developed in *M. galloprovincialis* in the last years. For instance, a comparative expression profile of mantle, muscle, gills and hemocytes was performed in 2015 by Moreira et al. [54]. This work allowed to define dominant biological processes of each tissue. Hemocytes and adductor muscle appeared more related to immune and defense processes, gills to recognition of non-self patterns and mantle to reproduction and shell formation. In a similar approach, Bjärnmark et al., 2016 [55] established by transcriptomics different groups of genes being expressed in three different sections of the mantle, essentially related to biomineralization. Also, the gene expression in different developmental stages, from oocytes to juveniles was studied in *M. galloprovincialis* [56]. In this work, apart from defining the signature profiles of each development stage, it could be stablished that oocytes which expressed a higher quantity of genes such as myticins were more likely to reach success in the offspring.

In a very recent work [39], it was performed a transcriptomic analysis of mussel gills after a bath infection with *Vibrio splendidus*. This study aimed to mimic a natural infection with potentially pathogenic bacteria for mussels. The researchers concluded that the main processes regulated in gills were related to recognition, activating effector agents of the immune response in order to overcome the bacterial infection.

The transcriptome encompasses coding and noncoding RNAs. Although traditionally protein-coding mRNAs have been the most studied, noncoding RNAs have gained importance in recent years. The most common technique currently used to carry out this type of studies is **RNA sequencing (RNA-Seq)**. This high-throughput sequencing method provide high coverage and resolution, providing an image of the phenotype of the organism. The advantages over previous methodologies are that the technique is based on the addition by a DNA polymerase of fluorescent nucleotides one by one onto a growing DNA template strand (in contrast to hybridization-based methods of microarrays) and that the method is massively parallel, sequencing millions of fragments simultaneously (while the Sanger method only

sequenced a single DNA fragment at a time), resulting in hundreds to thousands of genes at one time [57].

An RNA-Seq experiment begins with the isolation of RNA. It is necessary that the RNA have a good quality, fact that is measured by the RNA Integrity Number (RIN), based on ribosomic RNAs. Once the RNA is suitable, the sequencing library has to be prepared. This part encompasses the retrotranscription of the RNA to cDNA, fragmentation and ligation of sequencing adaptors. After that, the sequencing is performed in a NGS platform. The dominant platform today is Illumina, in which the DNA molecules are clonally amplified while immobilized on the surface of a glass flowcell [58].

No consesus pipeline has been established for the analysis of the RNA-Seq results. It is essentially subjected to the species of work and the previous available genomic information. In *M. galloprovincialis*, the use of the genome as a reference for mapping reads is complicated because of the presence/absence variation singularity. For this reason, the analysis protocol must be adjusted to the characteristics of each experiment. The method of analysis used in this thesis consists in the **trimming** of raw reads to remove low quality sequences, adaptor sequences, and short sequences. Then, the trimmed reads of all the samples are **assembled** in a reference global transcriptome. After that, an **RNA-Seq analysis** is performed, obtaining the expression values of each contig. Finally, a **differential expression analysis** test (the test used throughout this thesis is a Robinson and Smyth's Exact Test, which assumes a Negative Binomial distribution of the data and takes into account the overdispersion caused by biological variability) is executed to compare expression levels in each sample and to find the differentially expressed genes (DEGs).

Several transcriptomic studies based on RNA-Seq methodology and using the Illumina HiSeq™ 4000 technology were developed throughout this thesis. The aforementioned analysis pipeline is shown in the figure 5.

New technologies have allowed to advance in knowledge, especially in non-model species as mussels. These animals have a great potential from a genomic point of view, given its great variability, as

well as from an immune point of view, since they are exclusively provided with innate immunity and yet they still have a great capacity to resist diseases and to colonize new territories. These features make the mussel a very interesting model species in a field of comparative immunology.

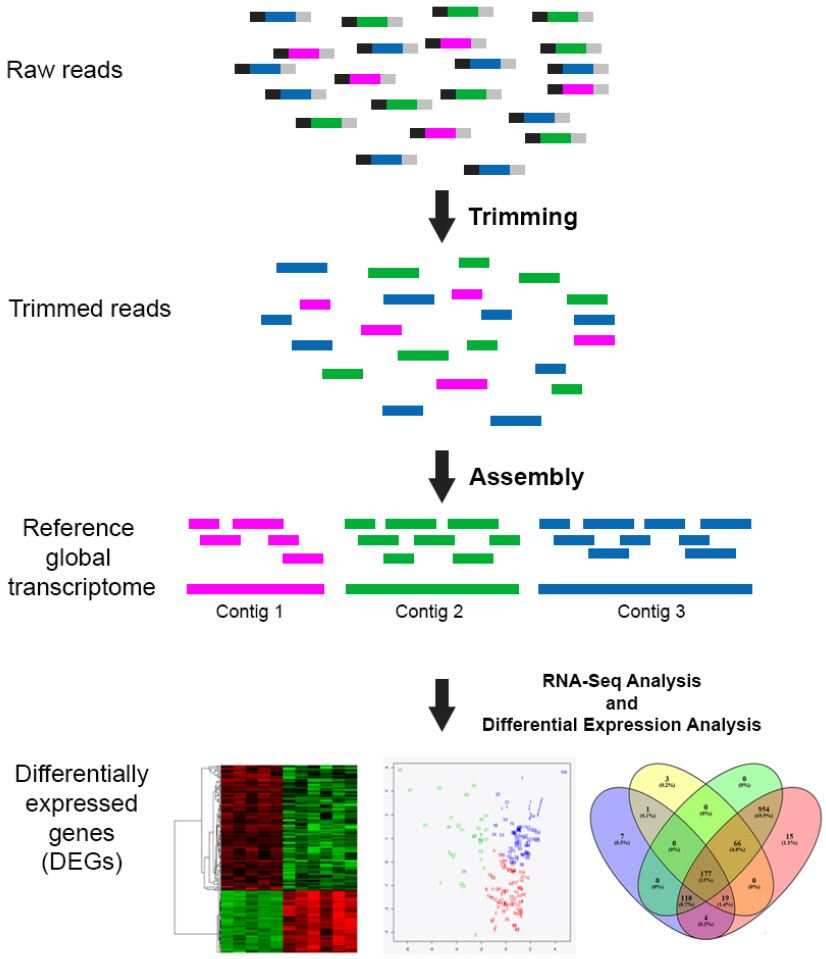

Figure 5. Schematic representation of the RNA-Seq pipeline. Figure source: own work.

1.1.6 References

1. United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019, custom data acquired via website (accessed on 19 August 2020).
2. FAO. FAO Home Page. Available online: http://www.fao.org/home/en/ (accessed on 17 August 2020).
3. Bonham, V. *Mytilus galloprovincialis*. Invasive Species Compendium (CAB, Wallingford, UK). (2017).
4. Goldberg, E.D. The mussel watch — A first step in global marine monitoring. Mar Pollut Bull. 6, 111 (1975).
5. Fiorito, F.; Amoroso, M.G.; Lambiase, S.; Serpe, F.P.; Bruno, T.; Scaramuzzo, A.; Maglio, P.; Fusco, G.; Esposito, M. A relationship between environmental pollutants and enteric viruses in mussels (*Mytilus galloprovincialis*). Environ Res. 169, 156–162 (2019). doi:10.1016/j.envres.2018.11.001.
6. Romero, A.; del Mar 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. 108, 149–163 (2014). doi:10.3354/dao02701.
7. Benabdelmouna, A.; Garcia, C.; Ledu, C.; Lamy, P.; Maurouard, E.; Dégremont, L. Mortality investigation of *Mytilus edulis* and *Mytilus galloprovincialis* in France: An experimental survey under laboratory conditions. Aquaculture. 495, 831–841 (2018). doi:10.1016/j.aquaculture.2018.06.075.
8. Garcia, C.; Haond, C.; Chollet, B.; Nerac, M.; Omnes, E.; Joly, J.P.; Dubreuil, C.; Serpin, D.; Langlade, A.; Le Gal, D.; Terre-Terrillon, A.; Courtois, O.; Guichard, B.; Arzul, I. Descriptions of *Mikrocytos veneroïdes* n. sp. and *Mikrocytos donaxi* n. sp. (Ascetosporea: Mikrocytida: Mikrocytiidae), detected during important mortality events of the wedge clam *Donax trunculus* Linnaeus (Veneroida:

Donacidae), in France between 2008 and 2011. Parasit Vectors. 11, 119 (2018). doi:10.1186/s13071-018-2692-0.

9. Segarra, A.; Pepin, J.F.; Arzul, I.; Morga, B.; Faury, N.; Renault, T. Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res. 153, 92–99 (2010). doi:10.1016/j.virusres.2010.07.011.
10. Medzhitov, R.; Janeway, C.A.Jr. Innate immunity: the virtues of a nonclonal system of recognition. Cell. 91, 295–298 (1997). doi:10.1016/s0092-8674(00)80412-2.
11. El Chamy, L.; Leclerc, V.; Caldelari, I.; Reichhart, J.M. Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. Nat Immunol. 9, 1165–1170 (2008). doi:10.1038/ni.1643.
12. Berisha, A.; Mukherjee, K.; Vilcinskas, A.; Spengler, B.; Römpp, A. High-resolution mass spectrometry driven discovery of peptidic danger signals in insect immunity. PLoS One. 8, e80406 (2013). doi:10.1371/journal.pone.0080406.
13. Allam, B.; Raftos, D. Immune responses to infectious diseases in bivalves. J Invertebr Pathol. 131, 121–36 (2015). doi:10.1016/j.jip.2015.05.005.
14. Costa, M.M.; Prado-Alvarez, M.; Gestal, C.; Li, H.; Roch, P.; 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. 26, 515–23 (2009). doi:10.1016/j.fsi.2009.02.001.
15. Gerdol, M.; Venier, P. An updated molecular basis for mussel immunity. Fish Shellfish Immunol. 46, 17–38 (2015). doi:10.1016/j.fsi.2015.02.013.
16. Venier, P., De Pittà, C., Bernante, F., Varotto, L., De Nardi, B., Bovo, G., Roch, P., Novoa, B., Figueras, A., Pallavicini, A., Lanfranchi, G. MytiBase: a knowledgebase of mussel (*M.*

galloprovincialis) transcribed sequences. BMC genomics. 10, 72 (2009). doi: 10.1186/1471-2164-10-72.

17. Gerdol, M.; Manfrin, C.; De Moro, G.; Figueras, A.; Novoa, B.; Venier, P.; Pallavicini, A. 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 (2011). doi:10.1016/j.dci.2011.01.018.

18. Gerdol, M.; Venier, P.; Pallavicini, A. The genome of the Pacific oyster *Crassostrea gigas* brings new insights on the massive expansion of the C1q gene family in Bivalvia. Dev Comp Immunol, 49, 59–71 (2015). doi:10.1016/j.dci.2014.11.007.

19. Li, D.; Nie, H.; Dong, S.; Huo, Z.; Yan, X. Molecular cloning and expression analysis of C-type lectin (RpCTL) in Manila clam *Ruditapes philippinarum* after lipopolysaccharide challenge. Fish Shellfish Immunol. 86, 981–993 (2019). doi:10.1016/j.fsi.2018.12.033.

20. Wang, W.; Gong, C.; Han, Z.; Lv, X.; Liu, S.; Wang, L.; Song, L. The lectin domain containing proteins with mucosal immunity and digestive functions in oyster *Crassostrea gigas*. Fish Shellfish Immunol. 89, 237–247 (2019). doi:10.1016/j.fsi.2019.03.067.

21. Hanington, P.C.; Zhang, S.M. The primary role of fibrinogen-related proteins in invertebrates is defense, not coagulation. J Innate Immun. 3, 17–27(2011). doi:10.1159/000321882.

22. Romero, A.; Dios, S.; Poisa-Beiro, L.; Costa, M.M.; 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. 35, 334–344 (2011). doi:10.1016/j.dci.2010.10.007.

23. Toubiana, M.; Gerdol, M.; Rosani, U.; Pallavicini, A.; Venier, P.; Roch, P. Toll-like receptors and MyD88 adaptors in Mytilus:

complete cds and gene expression levels. Dev Comp Immunol. 40, 158–66 (2013). doi:10.1016/j.dci.2013.02.006.

24. Petri, B.; Sanz, M.J. Neutrophil chemotaxis. Cell Tissue Res. 371, 425–436 (2018). doi:10.1007/s00441-017-2776-8.

25. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS One. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

26. Gordon S. Phagocytosis: An Immunobiologic Process. Immunity. 44, 463–475 (2016). doi:10.1016/j.immuni.2016.02.026.

27. Arumugam, M.; Romestand, B.; Torreilles, J.; Roch, P. *In vitro* production of superoxide and nitric oxide (as nitrite and nitrate) by *Mytilus galloprovincialis* haemocytes upon incubation with PMA or laminarin or during yeast phagocytosis. Eur J Cell Biol. 79, 513–519 (2000). doi:10.1078/0171-9335-00068.

28. Parisi, M.G.; Maisano, M.; Cappello, T.; Oliva, S.; Mauceri, A.; Toubiana, M.; Cammarata, M. Responses of marine mussel *Mytilus galloprovincialis* (Bivalvia: Mytilidae) after infection with the pathogen *Vibrio splendidus*. Comp Biochem Physiol C Toxicol Pharmacol. 221, 1–9 (2019). doi:10.1016/j.cbpc.2019.03.005.

29. Sendra, M.; Saco, A.; Yeste, M.P.; Romero, A.; Novoa, B.; Figueras, A. Nanoplastics: From tissue accumulation to cell translocation into *Mytilus galloprovincialis* hemocytes. resilience of immune cells exposed to nanoplastics and nanoplastics plus *Vibrio splendidus* combination. J Hazard Mater. 388, 121788 (2020). doi:10.1016/j.jhazmat.2019.121788.

30. Hubert, F.; Noel, T.; Roch, P. A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem. 240, 302–306 (1996). doi:10.1111/j.1432-1033.1996.0302h.x.

31. Mitta, G.; Hubert, F.; Dyrynda, E.A.; Boudry, P.; Roch, P. Mytilin B and MGD2, two antimicrobial peptides of marine mussels: gene

structure and expression analysis. Dev Comp Immunol. 24, 381–393 (2000). doi:10.1016/s0145-305x(99)00084-1.

32. Sonthi, M.; Toubiana, M.; Pallavicini, A.; Venier, P.; Roch, P. Diversity of coding sequences and gene structures of the antifungal peptide mytimycin (MytM) from the Mediterranean mussel, *Mytilus galloprovincialis*. Mar Biotechnol. 13, 857–867 (2011). doi:10.1007/s10126-010-9345-4.

33. Mitta, G.; Hubert, F.; Noël, T.; Roch, P. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur J Biochem. 265, 71–78 (1999). doi:10.1046/j.1432-1327.1999.00654.x.

34. Pallavicini, A.; Costa M.M.; Gestal, C.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

35. Costa, M.M.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

36. Sonthi, M.; Cantet, F.; Toubiana, M.; Trapani, M.R.; Parisi, M.G.; Cammarata, M.; Roch, P. Gene expression specificity of the mussel antifungal mytimycin (MytM). Fish Shellfish Immunol. 32, 45–50 (2012). doi:10.1016/j.fsi.2011.10.017.

37. Li, H.; Parisi, M.G.; Toubiana, M.; Cammarata, M.; Roch, P. Lysozyme gene expression and hemocyte behaviour in the Mediterranean mussel, *Mytilus galloprovincialis*, after injection of various bacteria or temperature stresses. Fish Shellfish Immunol. 25, 143–152 (2008). doi:10.1016/j.fsi.2008.04.001.

38. Rosani, U.; Varotto, L.; Gerdol, M.; Pallavicini, A.; Venier, P. IL-17 signaling components in bivalves: Comparative sequence analysis and involvement in the immune responses. Dev Comp Immunol. 52, 255–268 (2015). doi:10.1016/j.dci.2015.05.001.

39. Saco, A.; Rey-Campos, M.; Novoa, B.; Figueras, A. Transcriptomic response of mussel gills against *Vibrio splendidus* reveals its role in the immune response. Front Immunol. 11, 3273 (2020). doi:10.3389/fimmu.2020.615580.

40. Rosani, U.; Varotto, L.; Domeneghetti, S.; Arcangeli, G.; Pallavicini, A.; Venier, P. Dual analysis of host and pathogen transcriptomes in ostreid herpesvirus 1-positive *Crassostrea gigas*. Environ Microbiol. 17, 4200–4212 (2015). doi:10.1111/1462-2920.12706.

41. Netea, M.G.; Quintin, J.; Van der Meer, J.W.M. Trained immunity: a memory for innate host defense. Cell Host Microbe. 9, 355–61 (2011). doi:10.1016/j.chom.2011.04.006.

42. Netea, M.G.; Joosten, L.A.; Latz, E.; Mills, K.H.; Natoli, G.; Stunnenberg, H.G.; O'Neill, L.A.; Xavier, R.J. Trained immunity: A program of innate immune memory in health and disease. Science. 352, aaf1098 (2016). doi:10.1126/science.aaf1098.

43. Gourbal, B.; Pinaud, S.; Beckers, G.J.M.; Van Der Meer, J.W.M.; Conrath, U.; Netea, M.G. Innate immune memory: An evolutionary perspective. Immunol Rev. 283, 21-40 (2018). doi: 10.1111/imr.12647.

44. Sharrock, J.; Sun, J.C. Innate immunological memory: from plants to animals. Curr Opin Immunol. 62, 69-78 (2020). doi: 10.1016/j.coi.2019.12.001.

45. Kurtz, J.; Franz, K. Evidence for memory in invertebrate immunity. Nature. 425, 37–38 (2003). https://doi.org/10.1038/425037a.

46. Sadd, B.M.; Schmid-Hempel, P. Insect immunity shows specificity in protection upon secondary pathogen exposure. Curr Biol. 20, 1206-1210 (2006). doi:10.1016/j.cub.2006.04.047.

47. Melillo, D.; Marino, R.; Italiani, P.; Boraschi, D. Innate Immune Memory in Invertebrate Metazoans: A Critical Appraisal. Front Immunol. 22, 1915 (2018). doi:10.3389/fimmu.2018.01915.

48. Green, T.J.; Montagnani, C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against

subsequent challenge with Ostreid herpesvirus (OsHV-1 µvar). Fish Shellfish Immunol. 35, 382–8 (2013). doi:10.1016/j.fsi.2013.04.051.

49. Lafont, M.; Petton, B.; Vergnes, A.; Pauletto, M.; Segarra, A.; Gourbal, B.; Montagnani, C. Long-lasting antiviral innate immune priming in the Lophotrochozoan Pacific oyster, *Crassostrea gigas*. Sci Rep. 7, 13143 (2017). doi:10.1038/s41598-017-13564-0.

50. Lafont, M.; Vergnes, A.; Vidal-Dupiol, J.; de Lorgeril, J.; Gueguen, Y.; Haffner, P.; Petton, B.; Chaparro, C.; Barrachina, C.; Destoumieux-Garzon, D.; Mitta, G.; Gourbal, B.; Montagnani, C. A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, *Crassostrea gigas*. mBio. 10, e02777-19 (2020). doi: 10.1128/mBio.02777-19.

51. Green, T.J.; Helbig, K.; Speck, P.; Raftos, D.A. Primed for success: oyster parents treated with poly(I:C) produce offspring with enhanced protection against Ostreid herpesvirus type I infection. Mol Immunol. 78, 113–20 (2016). doi:10.1016/j.molimm.2016.09.002.

52. Green, T.J.; Speck, P. Antiviral Defense and Innate Immune Memory in the Oyster. Viruses. 10, 133 (2018). doi: 10.3390/v10030133.

53. Gerdol, M.; Moreira, R.; Cruz, F.; Gómez-Garrido, J.; Vlasova, A.; Rosani, U.; Venier, P.; Naranjo-Ortiz, M.A.; Murgarella, M.; Greco, S.; Balseiro, P.; Corvelo, A.; Frias, L.; Gut, M.; Gabaldón, T.; Pallavicini, A.; Canchaya, C.; Novoa, B.; Alioto, T.S.; Posada, D.; Figueras, A. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome Biol. 21, 275 (2020). doi:10.1186/s13059-020-02180-3.

54. Moreira, R.; Pereiro, P.; Canchaya, C.; Posada, D.; Figueras, A.; Novoa, B. RNA-Seq in *Mytilus galloprovincialis*: comparative transcriptomics and expression profiles among different tissues. BMC Genomics. 16, 728 (2015). doi:10.1186/s12864-015-1817-5.

55. Bjärnmark, N.A.; Yarra, T.; Churcher, A.M.; Felix, R.C.; Clark, M.S.; Power, D.M. Transcriptomics provides insight into *Mytilus galloprovincialis* (Mollusca: Bivalvia) mantle function and its role in biomineralisation. Mar genomics. 27, 37–45 (2016). doi:10.1016/j.margen.2016.03.004.

56. Moreira, R.; Pereiro, P.; Balseiro, P.; Milan, M.; Pauletto, M.; Bargelloni, L.; Novoa, B.; Figueras, A. Revealing *Mytilus galloprovincialis* transcriptomic profiles during ontogeny. Dev Comp Immunol. 84, 292–306 (2018). doi:10.1016/j.dci.2018.01.016.

57. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 10, 57–63 (2009). doi:10.1038/nrg2484.

58. Kukurba, K.R.; Montgomery, S.B. RNA Sequencing and Analysis. Cold Spring Harb Protoc. 2015, 951–969 (2015). doi:10.1101/pdb.top084970.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1.2 OBJECTIVES

The main objective of this thesis was to increase the knowledge of the immune response of mussels, especially genetic, transcriptomic and functional variability of genes such as the antimicrobial peptides (AMPs).

In particular:

1. Increase the transcriptomic information about how mussel hemocyte respond after a *Vibrio splendidus* infection, paying particular attention to the individual response.
2. Study the transcriptomic response of mussels hemocytes after two exposures to the same pathogen (*Vibrio splendidus*). The interest of this study lies in the recently reported memory capacity of the innate immune system in invertebrates. It is important to note that the different mussels were sampled as naïve mussels and after the two successive infections, getting an idea of how the response evolved over the time.
3. Analyze the transcriptomic response of the hemocytes after being stimulated with synthetic myticin C. These peptides are among the most expressed genes of mussel hemocytes and they have classically demonstrated antimicrobial functions. However, some studies suggested that they could act as a chemokine-like molecule. Because of that, we tried to study more in detail this new function of myticins.
4. Study the sequence of the myticin gene, trying to find the origin of the variability as well as evolutionary information about this family of antimicrobial peptides that are key for mussels and their immune response.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 2

HIGH INDIVIDUAL VARIABILITY IN THE TRANSCRIPTOMIC RESPONSE OF MEDITERRANEAN MUSSELS TO *VIBRIO* REVEALS THE INVOLVEMENT OF MYTICINS IN TISSUE INJURY

Rey-Campos, M.[1]; Moreira, R.[1]; Valenzuela-Muñoz, V.[2]; Gallardo-Escárate, C.[2]; Novoa, B.[1]; Figueras, A.[1]. High individual variability in the transcriptomic response of Mediterranean mussels to *Vibrio* reveals the involvement of myticins in tissue injury. Sci Rep. 9, 3569 (2019). doi:10.1038/s41598-019-39870-3. (Open Access).

[1]Institute of Marine Research (IIM), National Research Council (CSIC), Eduardo Cabello, 6, 36208, Vigo, Spain. [2]Laboratory of Biotechnology and Aquatic Genomics, Interdisciplinary Center for Aquaculture Research (INC AR), University of Concepción, P.O. Box 160-C, Concepción, Chile.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 2: High individual variability in the transcriptomic response of mediterranean mussels to *Vibrio* reveals the involvement of myticins in tissue injury

2.1 INTRODUCTION

In the past years, a considerable effort has been made to understand the molecular basis of many biological processes of non-model organisms, such as bivalves. Although they are cultured worldwide and have an important ecological value, we are still far from understanding how these animals respond to pathogens. Mussels (*M. galloprovincialis*), according to their way of life, based on a filtration feed, can filter on average 7.5 liters of water in one hour [1], meaning they are in intimate contact with millions of microorganisms that are potentially pathogenic to them [2]. In fact, Stabili et al. [3] reported that the abundance of *Vibrio spp.* is higher in mussels than in the surrounding water. This constant challenge can compromise the health status of the bivalves and affect their culture, producing economic losses all over the world. As an example, the ostreid herpesvirus 1 (OsHV-1) caused massive mortalities in oysters (*Crassostrea gigas*) in different parts of the world [4,5]. In a similar way, in 2010 and 2011, strong increases in mortality were reported in different wild beds of the wedge clam *Donax trunculus*. This research work indicated that the pathogens responsible were parasites belonging to the genus *Mikrocytos* [6]. However, massive mortalities have never been reported in field for *M.*

galloprovincialis [7,8], despite cohabitate in the same areas as oysters and clams.

Instead lacking adaptive immunity, mussels recognize pathogens through their PAMPs [9,10], but also danger signals, and danger/damage-associated molecular patterns (DAMPs) [11,12]. The way to respond to these PAMPs, and even more to DAMPs, remains poorly studied. To understand the immune reaction of these animals, a significant effort has been made in recent years to increase the genomic and transcriptomic resources in bivalves and specifically in the Mediterranean mussel, *M. galloprovincialis* [13–18].

To date, all the genomic studies carried out in mussels have been conducted using biological pooled samples disregarding the importance of analyzing the individual response. However, it is known that mussels have immune effectors that are tremendously variable within the population [19–21]. This is the case of myticins, which are antimicrobial peptides highly expressed in mussel hemocytes [22]. Three different myticin genes have been defined and named as A, B and C so far, with similar DNA sequences and physicochemical properties [22,23]. Of them, myticin C is the most variable AMP in mussels with broad biological properties [24].

In this work, we analyzed the transcriptomic response of six individual naïve mussels and how they respond to a simple injury (control animals injected with filtered sea water) or against a bacterial challenge (animals infected with *Vibrio splendidus*). Our results highlight the variability of the response in individual mussels and the importance of the appropriate experimental controls since the injection on the adductor muscle (the usual method to experimentally infect mussels) might be interpreted by the organism as a danger signal that may influence the unexpected regulation of several expressed genes.

2.2 MATERIALS AND METHODS

2.2.1 Animals

Adult *M. galloprovincialis*, 6–8 cm in shell length, were obtained from a commercial shellfish farm (Vigo, Galicia, Spain) and maintained

in open-circuit filtered sea water tanks at 15°C with aeration. The animals were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, the animals were acclimatized to aquarium conditions for at least one week.

2.2.2 Experimental design

A schematic representation of the experimental design is shown in figure 1.

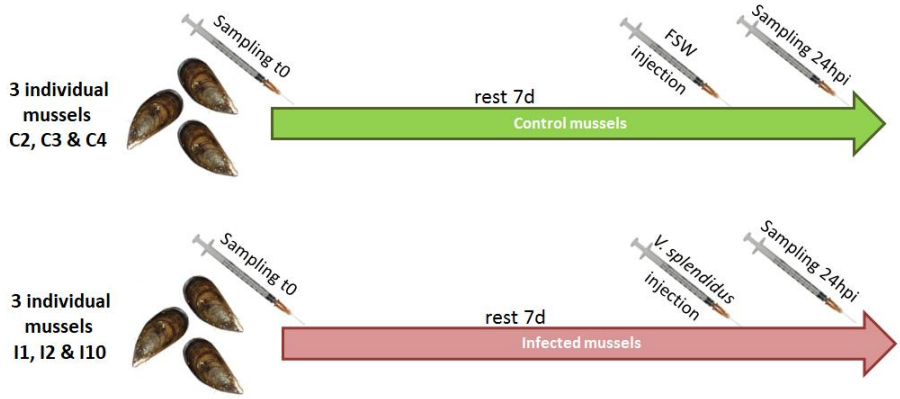

Figure 1. Diagram of the experimental design. Three control mussels (C2, C3 and C4) and three infected mussels (I1, I2 and I10) were selected for sequencing.

Twenty naïve mussels were marked and notched on the shell, and hemolymph (500 µl) was withdrawn from the adductor muscle of each mussel with a 0.5 mm diameter (25 G) disposable needle. This sampling point corresponds to time zero (t0). The hemolymph was centrifuged at 4°C at 3,000 g for 10 min, and the pellet was resuspended in 500 µl of TRIzol (Invitrogen), immediately homogenized with syringe and a 0.5 mm diameter (25 G) disposable needle and stored at −80°C until RNA isolation. After one week, 10 of the mussels were injected in the adductor muscle with 100 µl of filtered sea water (FSW). The other 10 mussels were injected in the same way with 100 µl of a solution of *Vibrio splendidus* (reference strain, LGP32) [7] at a non-lethal

concentration (1×10^7 CFU/mL). One day after the challenge, that is, 24 hours post injection (24 hpi), hemolymph (500 µl) was sampled again from individual mussels, centrifuged in the same manner described above, and the pellet was resuspended in 500 µl of TRIzol. Samples were immediately homogenized with syringe and 25 G needle and kept at −80°C until RNA isolation.

2.2.3 RNA isolation, cDNA production and Illumina sequencing

RNA isolation was carried out in the 40 samples (n=20 naïve at t0, n=10 FSW injected at 24hpi, and n=10 bacteria injected at 24hpi) using TRIzol and following the manufacturer's protocol. Purification of RNA after DNase I treatment was performed with RNeasy mini (Qiagen). Next, the concentration and purity of the RNA was measured using a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Inc.), and RNA integrity was tested on an Agilent 2100 Bioanalyzer (Agilent Technologies) before producing cDNA libraries for Illumina sequencing. Only the individuals with the best RNA samples (in terms of RNA quantity and quality) from both sampling points were chosen for Illumina sequencing: control n°2 (C2), control n°3 (C3) control n°4 (C4), infected n°1 (I1), infected n°2 (I2) and infected n°10 (I10). In total, 12 RNA samples (2 per individual, the first at t0 and the second 24hpi of FSW or bacteria) were sequenced (details in Table 1).

The mRNA-Seq sample preparation kit from Illumina was used according to the manufacturer's instructions. 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 Qubit system. Next-generation sequencing was performed using Illumina HiSeq™ 4000 technology at Macrogen Inc. Korea (Seoul, Republic of Korea). The raw sequencing data have been deposited in

the NCBI Short Read Archive database under the accession ID SRP145077.

Table 1. Summary of transcriptome bioinformatics details.

Reads origin	Raw reads	Trimmed reads
C2 t0	78,426,948	99.59%
C3 t0	44,346,854	98.03%
C4 t0	100,814,198	99.59%
I1 t0	17,696,894	98.08%
I2 t0	93,114,098	99.60%
I10 t0	96,780,602	99.64%
C2 24h	95,296,484	99.49%
C3 24h	51,708,988	97.62%
C4 24h	92,661,282	99.65%
I1 24h	52,102,302	98.77%
I2 24h	90,965,875	99.52%
I10 24h	99,262,970	99.55%
Assembly		
Contigs	270,324	
Range contig length	200-15,624	
Average contig length	512	
N50	574	
Blast		
Contigs identified by UniProt/Swiss-Prot	24.97%	
Contigs identified by molluscs database	54.93%	
GO analysis		
Annotated contigs	24.87%	
KEGG analysis		
Pathways assigned to contigs	8.03%	

2.2.4 Bioinformatics and RNA-Seq

CLC Genomics Workbench, v.10.0.1 (CLC Bio; Qiagen) was used to filter, assemble and perform the RNA-Seq and the statistical analysis of individual mussels. Raw reads were trimmed to remove low quality sequences (PHRED = 13), adaptor sequences, and sequences shorter than 70 bp. Then, a reference global transcriptome of the six mussels was assembled with an overlap criterion of 70% and a similarity of 90% to exclude paralogous sequence variants. The settings used were a

mismatch cost = 2, deletion cost = 3, insert cost = 3, and minimum contig length = 200 base pairs. Before the expression analysis, a subsampling step of the trimmed reads was carried out. A random subset of sequences was generated to equal the number of reads present in each sample. Then, RNA-Seq analysis of the subsamples (mismatches = 2, length fraction = 0.8, similarity fraction = 0.8, and maximum hits per read = 10) was performed. The expression values were set as transcripts per million (TPM). Finally, a differential expression analysis test (a Robinson and Smyth's Exact Test, which assumes a Negative Binomial distribution of the data and takes into account the overdispersion caused by biological variability) was used to compare expression levels in each sample and to find the differentially expressed genes (DEGs). Transcripts with absolute fold change (FC) values > 2 and Bonferroni corrected p-value < 0.05 were retained for further analyses.

2.2.5 BLAST annotation, GO assignment, enrichment and KEGG analysis

UniProt/Swiss-Prot BLASTx results were used to obtain the Gene Ontology (GO) term assignments of the contig list using the Blast2GO software [25]. To improve the percentage of sequence identification, an in-house built database made with all of the mollusc sequences present in the NBCI nucleotide database was also used to perform the annotation. In both blast approaches, the e-value threshold was set at 1e-3. Then, the enrichment analysis of the up- and down-regulated DEGs (test set) were conducted, including the global mussel transcriptome as the reference set. A Fisher's exact test was run with default values and a p-value cut-off of 0.05. Only over-represented biological process (BP) terms were further analyzed. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in which DEGs were involved were also analyzed using the Blast2GO software and summarized following the existent categories in the KEGG database (http://www.genome.jp/kegg/pathway.html).

2.2.6 Flow cytometry analysis

The distribution of myticin C in different cell populations was assayed in mussel hemocytes by immunocytochemistry and flow cytometry (FACSCalibur; BD). The experimental design is depicted in figure 2.

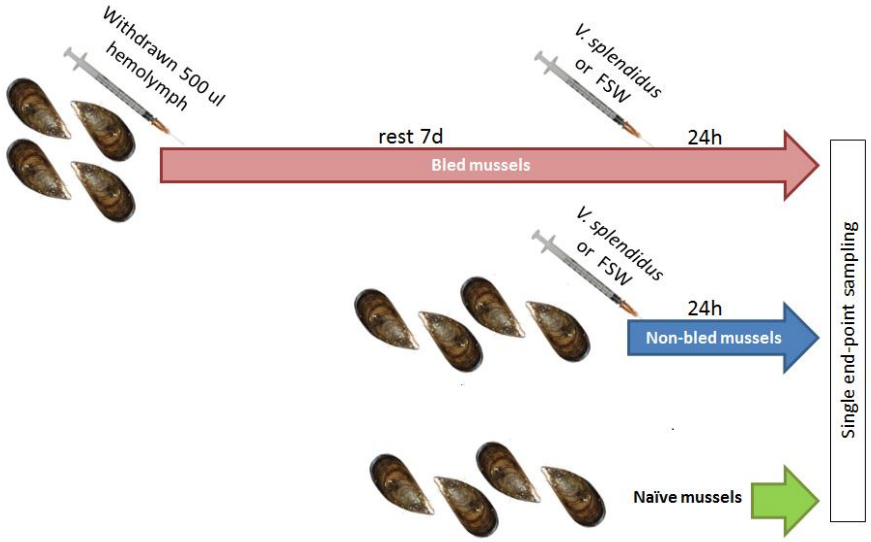

Figure 2. Experimental design for immunocytochemistry FACS detection of myticin C protein in mussel hemocytes. Four individual mussels were used for each experimental condition (naïve, non-bled FSW, non-bled *Vibrio*, bled FSW and bled *Vibrio*).

Briefly, the studied conditions were five: (1) naïve mussels, (2) mussels injected with FSW, (3) mussels injected with *V. splendidus*, (4) bled mussels and injected with FSW after a resting of 1 week and (5) bled mussels and injected with *V. splendidus* after a resting of 1 week. A single end-point sampling was performed for each condition. Four mussels per condition were analyzed. For each condition, 1.5 mL of hemolymph from individual mussels was withdrawn and immediately fixed in a final concentration of 2% paraformaldehyde (PFA). Samples were divided into three aliquots; the first one was used to analyze

myticin C production through a custom antibody [24]; the second one was exposed to the prebleed serum to be used as a control, and the third one was an absolute control without primary antibody. The immunocytochemistry protocol was performed as previously described in Balseiro et al. [24]. The cell suspension was fixed for 15 min at 4°C and washed twice for 10 min in phosphate buffered saline buffer (PBS) prior to permeabilization and staining with the anti-myticin C antibody (1:100). After overnight staining, samples were washed once for 10 min (PBS, 0.1% saponin, 0.2% bovine serum albumin, BSA) and incubated in dark conditions for 50 min at room temperature with secondary antibody (1:500). Finally, samples were washed for 10 min (PBS), and 200 µl of each condition and replicate were dispensed in a 96-well plate to be analyzed by fluorescence-activated cell sorting (FACS). The density plots and histograms were generated using Cell Quest Pro software (BD), and a one-way ANOVA with post-hoc analysis was used to analyze the significance of the results among all the conditions tested.

2.3 RESULTS

2.3.1 Assembly and annotation of mussel transcriptome

A summary of the sequence origin, assembly, identification, and annotation results is shown in Table 1. An average of 76 million raw reads was obtained from each individual sample of *M. galloprovincialis* hemocytes. The CLC Genomics Workbench was used to filter the raw reads, and over 97% of raw reads successfully passed the quality control in all the samples. The assembly step was performed with all the available samples to obtain a global mussel transcriptome, and 270,324 contigs were assembled with an average length of 512 bp. The putative identities of these sequences were obtained by Blast by two different ways; Blast2GO software was used to identify the 24.97% of the contigs through a BLASTx approach against UniProt/Swiss-Prot, and CLC Genomics Workbench was used to identify 54.93% of the contigs using an in-house designed database with all the sequences available in NCBI (National Center for Biotechnology Information) for molluscs.

GO terms were assigned to 24.87% of the contigs and enzyme codes to find KEGG pathways to 8.03% of the sequences.

2.3.2 Mussel transcriptome after bacterial or DAMP stimulation

The experimental design allowed us to sample hemolymph from each individual mussel before and after the injection with bacteria or FSW; therefore, the real behavior of the modulated genes could be followed in each animal. Figure 3 shows the distribution of the differentially expressed genes (DEGs) in control and infected animals 24 hpi compare to their own t0 sampling point (the whole set of genes, including the fold changes and annotation information, is available in the following link: https://static-content.springer.com/esm/art%3A10.1038%2Fs41598-019-39870-3/MediaObjects/41598_2019_39870_MOESM1_ESM.xlsx).

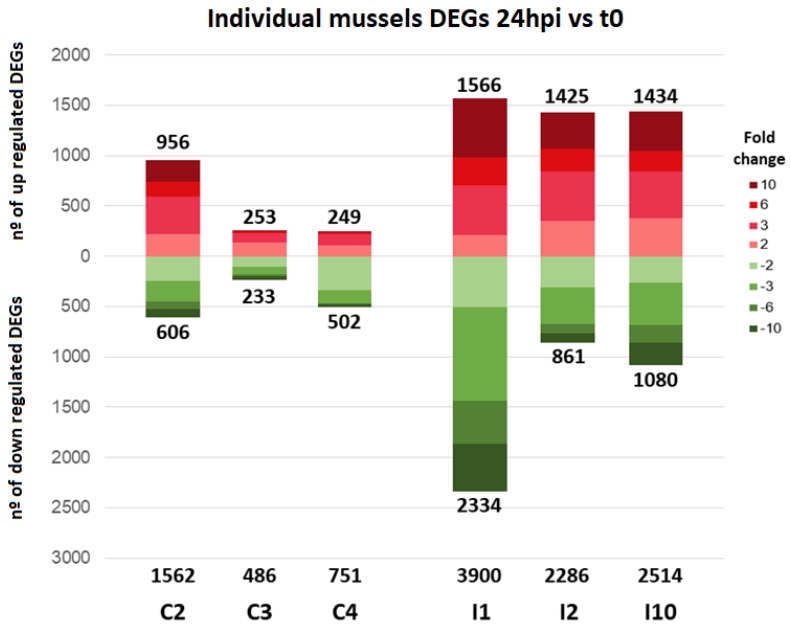

Figure 3. General overview of the expression values of mussels after bacterial challenge (infected) or FSW injection (controls). Stacked column charts reflect the fold change distribution of DEGs in control and infected animals 24 hpi with regard to their own t0 sampling point.

An expected response with more DEGs in infected animals (I1: 3,900 DEGs; I2: 2,286 DEGs; I10: 2,514 DEGs) compared to control animals (C2: 1,562 DEGs; C3: 486 DEGs; C4: 751 DEGs) was found. But, a very small percentage of modulated genes were shared between the 3 controls (only 2.9%) and the 3 infected animals (only 9.4%) as the Venn diagrams show (Figure 4a and 4b).

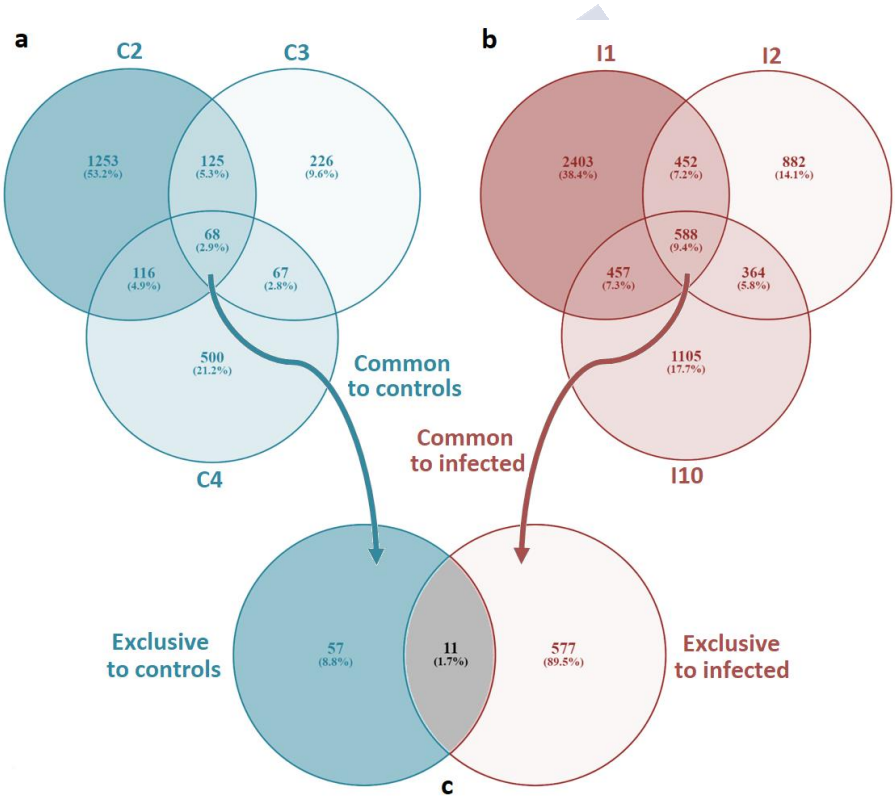

Figure 4. Venn diagrams of DEGs for each condition. **(a)** in blue, control mussels. **(b)** in red, infected mussels. **(c)** in gray, common genes to control and infected mussels.

This result could mean that each mussel responds to the same stimulus in a different way. However, a common response to each stimulus would be represented by the genes shared between the 3 individuals that constitute the group (FSW or *Vibrio splendidus*). Otherwise, comparing the 68 common genes to the 3 control animals and the 588 genes shared by the 3 infected individuals, 577 were exclusively modulated by the infection and 57 by the tissue damage produced by the FSW injection. This information is shown in figure 4c, as well as the number of common genes between both experimental conditions. Only 11 genes were shared between all the studied animals. Five out of 11 had informative annotations: C1q domain containing protein MgC1q61 (related to recognition in molluscs), chromobox protein homolog 7 (CBX7; epigenetic functions), putative gastrointestinal growth factor xP4 (important in mucosal protection and reconstitution), transcription factor AP-1 (a key transcription factor for immunity) and heat shock protein beta-1 (it shows numerous biological roles including regulation of stress resistance, inflammation and apoptosis). All these genes, regulated both in control and infected mussels, are directly or secondarily related to the defense response.

The enrichment analysis of the GO terms in each compartment of the Venn diagram is shown in figure 5, representing the 25 most significant BP in each group, divided into up- or down-regulated according to the proportion Test/Reference that they represent.

As far as the exclusive DEGs of control mussels are concerned (Figure 5a), unexpected results were found, with BPs related to the immune response, antimicrobial peptides, and more specifically, to myticins, such as “cell killing”. Moreover, some other processes related to immunity such as “regulation of lymphocyte mediated immunity”, “protection from natural killer cell mediated cytotoxicity” and “regulation of programmed cell death” appeared enriched in the analysis. Regarding DEGs exclusive to infected mussels (Figure 5b), the most interesting processes were linked to response to hypoxia, glucocorticoid, lipopolysaccharide and those related to cell proliferation, cell cycle or leucocyte activation, which may indicate an active state of hemocytes.

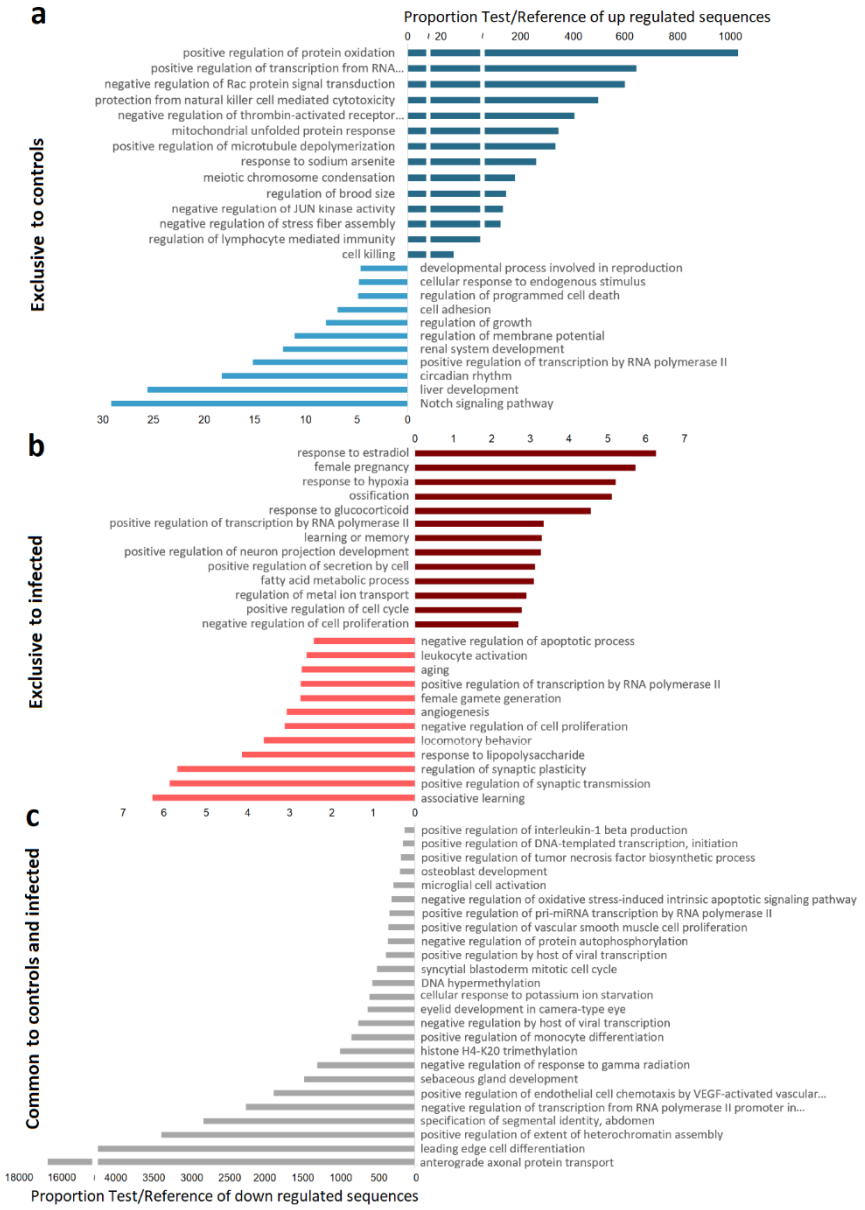

Figure 5. Enrichment analysis of DEGs. The proportions of test/reference sequences for up- and down- regulated genes are described following this legend:

(a) blue, exclusive genes for control mussels; **(b)** red, exclusive genes for infected mussels; **(c)** grey, common genes to control and infected mussels.

As stated above, the enrichment analysis revealed that the differentially BP represented in the common DEGs between control and infected mussels (Figure 5c) were related to immune terms such as "positive regulation of endothelial cell chemotaxis" representing the heat shock protein 27, "positive regulation of monocyte differentiation" or "negative regulation by host of viral transcription", both terms related to the transcription factor AP-1. GO terms related to histone or DNA methylation were also present: "histone H4-K20 trimethylation" and "DNA hypermethylation", representing the chromobox protein homolog 7 (CBX7). This result could indicate that both control and infected animals were subjected to a stimulus with similar capability to trigger epigenetic changes to respond to damage or infection.

The KEGG reference pathway analysis of control and infected DEGs were proportionally calculated (the percentage of the DEGs ascribed to every pathway with regard to the modulated genes in each Venn category). This analysis is summarized in figure 6, which shows that metabolism was affected in different ways in control and infected animals.

The most represented pathways in control mussels were related to nucleotide and carbohydrate metabolism as well as to the metabolism of cofactors and vitamins. In infected mussels, a significant number of DEGs were included in pathways known to be greatly affected by immune challenges such as signal transduction (phosphatidylinositol signaling system, the mTOR signaling pathway and the PI3K-Akt signaling pathway), or viral infectious diseases and immune system (defense cells differentiation and signaling). The PI3K, Akt and mTOR pathways are linked to the JAK-STAT signaling pathway, and their regulation is intimately related to the immune response and the regulation of processes such as cell proliferation, autophagy and apoptosis [26], in line with the enrichment analysis results.

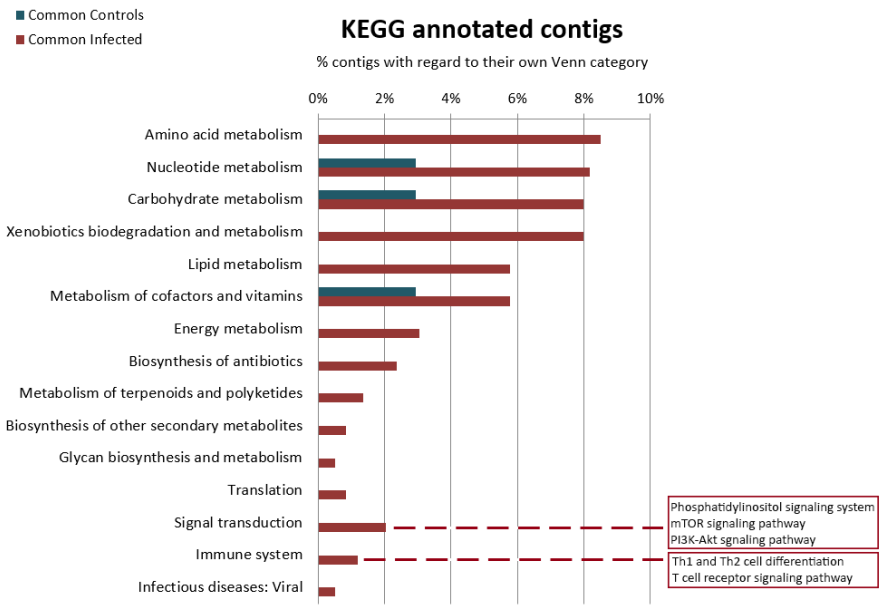

Figure 6. Summary of the KEGG reference pathway results for the significantly regulated contigs: common for control mussels (blue) and common for infected mussels (red).

With the aim of knowing more about the nature of these DEGs, we focused on the most expressed genes in each group. Tables 2, 3 and 4 show, respectively, the top 25 genes found exclusively in controls, in infected animals and all those regulated genes shared between both of them, with the fold changes for each individual mussel.

It is worth mentioning that antimicrobial peptides (AMPs) such as myticins or myticusin were highly up-regulated in FSW injected animals (Table 2) and not in the infected animals, most likely reflecting a reaction against an injury or a non-pathogenic danger signal. Interestingly, in the DEGs found only in controls, genes related to cell proliferation, differentiation or cell activation were also up-regulated and include elastin microfibril interface-located protein 2 (EMILIN-2), stathmin, low affinity epsilon Fc receptor, cell division cycle-associated protein 2 (CDCA2), and signal transducer and transcription activator (STAT). Some of these are associated with the cytoskeleton and cell

motility (for example stathmin or myticins that increase chemotaxis). The most down-regulated genes in controls were related to the regulation of cell death: cell death specification protein 2 (ces-2), immediate early response gene 5 protein (IER5), protocadherin Fat 1, and GTPase IMAP family member 4 (GIMAP4).

Table 2. Top 25 regulated DEGs associated to tissue injury (controls).

Contig	Description	Fold change			Mean
		C2	C3	C4	
Mg_109517	Myticin C	30.24	2.21	5.25	12.57
Mg_7618	Myticin B	25.00	2.09	2.85	9.98
Mg_127788	Myticin C	21.31	2.67	2.40	8.79
Mg_1042	EMILIN2	20.09	2.37	3.26	8.57
Mg_33725	Myticin C	15.24	2.99	2.08	6.77
Mg_5071	Stathmin	9.22	5.35	3.99	6.19
Mg_7746	Myticusin-alpha precursor	6.54	7.31	4.20	6.02
Mg_150217	Myticusin-alpha precursor	6.40	6.71	4.22	5.78
Mg_37432	Heavy metal-binding protein	8.39	3.62	4.20	5.40
Mg_26476	FCER2	7.49	3.56	2.71	4.59
Mg_27764	CDCA2	6.83	4.22	2.56	4.54
Mg_3246	Serine protease inhibitor A3C	7.60	3.79	2.15	4.51
Mg_5822	STAT	3.89	2.45	2.98	3.11
Mg_4540	STAT2	4.12	2.43	2.33	2.96
Mg_31595	SMC4	3.06	2.89	2.82	2.93
Mg_1195	NPAS4	-3.25	-3.09	-3.92	-3.42
Mg_1629	SREBF1	-3.81	-5.01	-2.54	-3.79
Mg_23409	Perlucin	-5.22	-3.84	-5.89	-4.98
Mg_20788	GIMAP4	-6.73	-2.92	-6.03	-5.23
Mg_157047	Protocadherin Fat 1	-5.63	-5.08	-5.87	-5.52
Mg_29303	NOTCH2NL	-5.48	-5.01	-6.44	-5.64
Mg_6692	IER5	-2.45	-12.17	-10.66	-8.43
Mg_10957	Ces-2	-3.92	-13.93	-8.16	-8.67
Mg_23430	Cornifelin homolog B	-15.91	-3.61	-34.40	-17.97
Mg_247	RING finger protein R811	-89.09	-3.30	-27.59	-39.99

In infected animals (Table 3), some of the exclusive DEGs were directly involved in the immune response, in concordance with the results of the KEGG pathways analysis (Figure 6).

Table 3. Top 25 regulated DEGs associated exclusively to a *Vibrio* infection.

Contig	Description	Fold change			Mean
		I1	I2	I10	
Mg_80300	POLR2A	441.85	89.21	1085.67	538.91
Mg_106	IRG1	655.70	221.92	370.69	416.10
Mg_457	SLC12A8	25.41	66.52	563.75	218.56
Mg_599	NPY2R	276.93	89.07	288.15	218.05
Mg_261923	Tropomyosin-1	26.44	69.72	153.20	83.12
Mg_3149	NPY2R	72.77	11.07	144.70	76.18
Mg_1646	Zasp52	50.97	117.84	32.44	67.08
Mg_784	Kpc-1	62.94	32.09	92.91	62.65
Mg_24087	SOCS2	164.09	17.37	4.69	62.05
Mg_32754	ZNF26	79.10	48.15	29.36	52.20
Mg_19011	PheS	42.23	47.93	58.54	49.57
Mg_2877	GHSR	92.45	38.46	6.31	45.74
Mg_9258	SOCS2	103.63	12.17	2.85	39.55
Mg_7966	PheS	43.11	28.66	43.79	38.52
Mg_6277	IFI44L	22.76	27.78	57.73	36.09
Mg_3834	MyD88	66.35	23.48	17.24	35.69
Mg_10058	PheS	37.23	23.16	46.01	35.46
Mg_6296	cGAS	64.94	17.83	12.45	31.74
Mg_17428	C1QL4	-40.12	-25.04	-40.69	-35.28
Mg_4119	COL14A1	-53.14	-30.65	-28.95	-37.58
Mg_21146	HSP	-53.09	-47.87	-25.07	-42.01
Mg_44317	Apextrin-like protein 1	-84.47	-8.82	-39.92	-44.40
Mg_3338	Headcase protein	-156.82	-16.21	-19.39	-64.14
Mg_11571	NDRG1	-59.97	-87.36	-75.87	-74.40
Mg_24372	C1QTNF4	-214.51	-18.16	-19.81	-84.16

Some of these genes that were up-regulated in the three tested animals were IRG1, SOCS2, IFI44L and Myd88. The cis-aconitate decarboxylase (also known as Immune-responsive gene 1 protein or IRG1) is an enzyme of the tricarboxylic acid cycle which produces itaconate after an immune challenge. Itaconate has been recently studied for its anti-inflammatory and antimicrobial properties [27,28], making a direct link between basic metabolism and the immune response. The suppressor of cytokine signaling 2 (SOCS2) is a well-known negative regulator of the JAK-STAT signaling cascade, whose function is to control the inflammatory response [29]. The myeloid differentiation primary response protein (Myd88), an essential mediator of the Toll signaling pathway, has been characterized in bivalves, and it is up-regulated after bacterial infections [30], consistent with our results. In contrast, the interferon-induced protein 44 (IFI44L) has been less studied, but the up-regulation of this gene has been associated with viral infections also in bivalves [31–33], and it is known to have antiviral properties [34]. However, the exact function is still unknown, and it could be possible that in bivalves, it also has antibacterial roles, such as its strong up-regulation 24 hours after *V. splendidus* injection suggested. The presence of IFI44L in the infected DEGs list is a possible explanation for the "viral infectious diseases" category in the KEGG analysis.

Table 4 shows the regulated transcripts shared between controls and infected animals. Two of them (C1q domain containing protein MgC1q61 and Putative gastrointestinal growth factor xP4) were up-regulated in all the controls but down-regulated in infected mussels. Individual responses can override experimental design in many genes, but in this case, even despite individual variability, it was possible to detect experimentally induced changes in hemocyte gene expression. In the specific case of the C1q domain containing protein, it could be due to its role as a pathogen recognition protein and its fast up-regulation after an infection, with a return to physiological levels in 24 h [21]. The putative gastrointestinal growth factor xP4 is a member of the trefoil factor family (TFF), a group of molecules with a pivotal role in maintaining the surface integrity of mucous epithelia *in vivo*, which explains its up-regulation after the injury in control mussels. Although

the exact role of the TFF peptides is unknown, they also induce antiapoptotic effects and probably modulate inflammatory processes [35,36], what could be a reason of its diverse behavior in control and infected samples. The other 3 annotated genes, always down-regulated, were the heat shock protein beta-1, associated with the acute phase response and found modulated after a tissue injury in *M. galloprovincialis* at the protein level [37]; chromobox protein homolog 7 (CBX7), related to epigenetics via chromatin remodeling and modification of histones [38]; and the transcription factor AP-1, known for controlling the expression of genes related to differentiation, proliferation and apoptosis, and intimately linked to NF-Kb [39].

Table 4. Regulated DEGs commons to controls and infected animals. FC: fold change.

Contig	Description	FC C2	FC C3	FC C4	Mean C	FC I1	FC I2	FC I10	Mean I
Mg_81902	C1Q61	7.96	8.54	6.85	7.78	-2.37	-3.77	-2.47	-2.87
Mg_257	xP4	3.67	2.28	2.27	2.74	-7.42	-2.16	-5.26	-4.95
Mg_384	CBX7	-2.34	-2.14	-3.09	-2.52	-2.84	-2.40	-3.48	-2.91
Mg_234	AP-1	-5.12	-4.36	-3.55	-4.34	-15.56	-8.14	-7.55	-10.42
Mg_496	HSPB1	-290.15	-25.12	-52.75	-122.67	-82.31	-94.11	-80.22	-85.55

2.3.3 Expression of myticins

The presence of myticins in the group of the most significantly up-regulated genes after FSW injection led us to further investigate this fact. Taking into account that we had previously found that these molecules had antibacterial and antiviral properties [24,40,41], it was expected to find them being up-regulated after bacterial challenge and not after a tissue injury, as they were in control animals. Figure 7 shows the expression of all the differentially expressed myticins in each individual. Control animals, after the injection of FSW, presented higher expression values than those challenged with *V. splendidus*, which show a down-regulation of myticins in one of the individuals (I2) or no regulation at all of the myticin transcripts in another individual (I10).

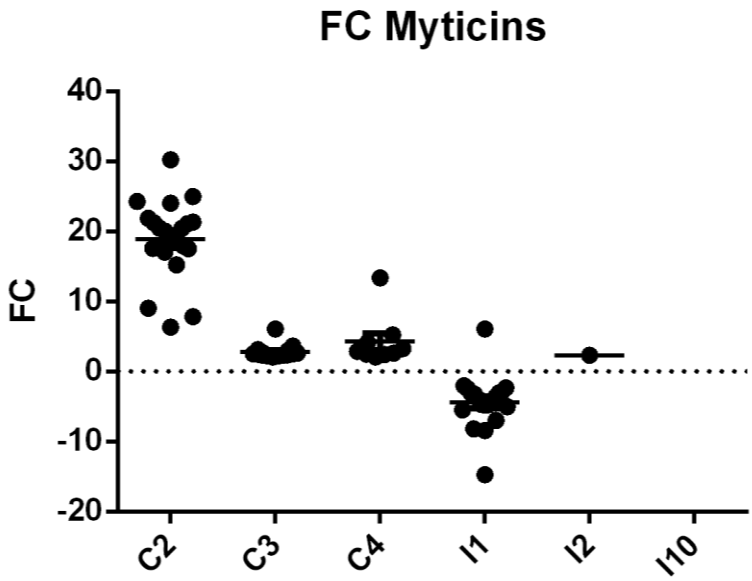

Figure 7. Expression profiles of myticins in control and infected individual mussels. Dots indicate the fold change (FC) of myticin transcripts 24 hpi compare to the t0. The line indicates the mean of the expression for each individual.

To further confirm these unexpected results, we designed a new experimental protocol to determine if myticins had the same behavior at the protein level using flow cytometry. The experiment had the same layout as the one used to study the transcriptome but with more experimental controls: we included naïve and non-bled mussels to confirm the involvement of tissue injury in the expression of myticins (Figure 2). Figure 8a shows the general profile of the hemocyte population, with R1 being the granulocytes and R2 the hyalinocytes. After *V. splendidus* infection in bled (injured) and non-bled mussels, there was a change in the hemocyte population structure with a significant reduction in the percentage of granulocytes with regard to their controls (FSW) and a significant increase in the percentage of hyalinocytes.

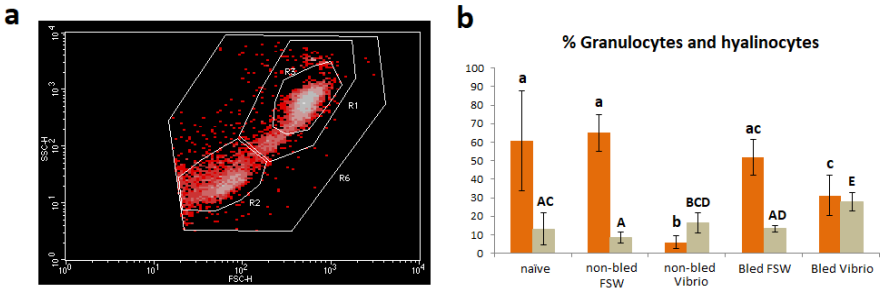

Figure 8. Description of hemocyte populations by flow cytometry. **(a)** Dot plot of the representative cell populations in mussel: R1, granulocytes; R2, hyalinocytes. Represented dots are gated for R6 region. **(b)** Description of the variation in hemocyte populations among samples: Orange, granulocytes; brown, hyalinocytes. Standard deviation is indicated for the four replicates. Different letters indicate significant differences (p-value < 0.05) among groups: low-case letters for granulocytes and capital letters for hyalinocytes.

These changes were present in bled and non-bled animals after infection, but the decrease in the granulocytes was less important in bled or injured mussels (Figure 8b). It is noteworthy that the hemocyte population changed drastically after bacterial injection, but it recovered significantly, with regard to the non-bled mussels, if a previous damage stimulus had been made (Figure 8b). Figure 9a shows the region considered positive for myticin C expression. These positive cells

corresponded to the granulocyte (R1) population (Figure 9b). When the production of myticin C was evaluated after the different treatments, a significant decrease in the positive for myticin C population was found after *Vibrio* infection (Figure 9c). However, this reduction was not observed when mussels were bled before, generating a tissue injury.

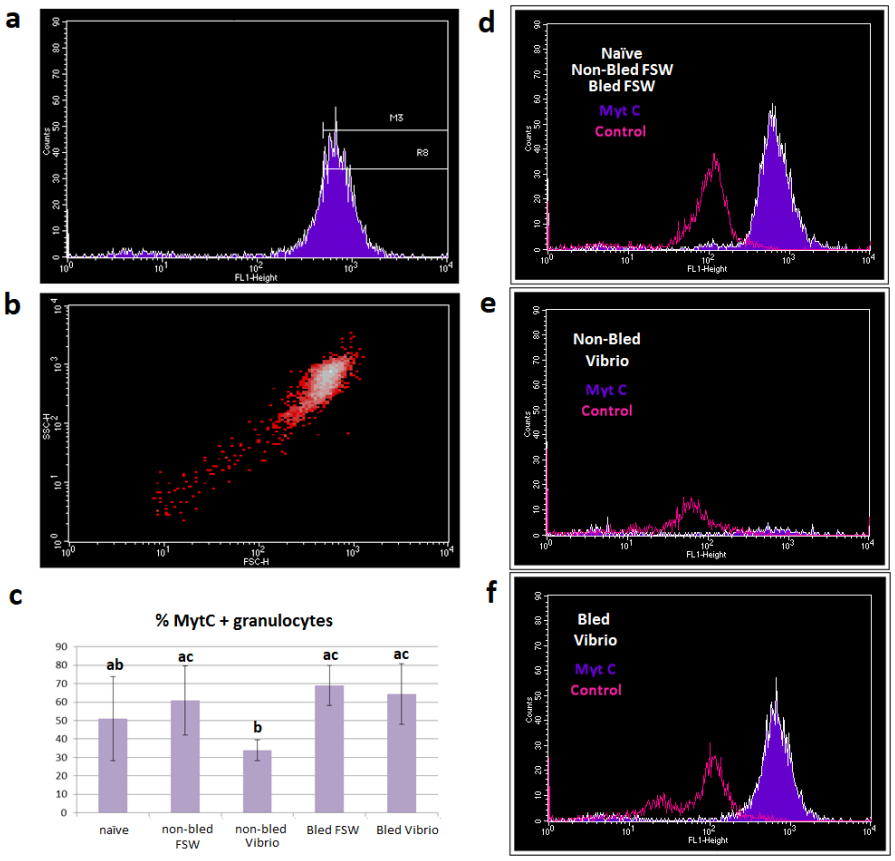

Figure 9. FACS results of the immunocytochemistry of myticin C. **(a)** Criteria to choose myticin C positive cells (M3). **(b)** Myticin C positive cells in a representative sample. **(c)** Statistical analysis of the percentage of granulocytes positive for myticin C immunocytochemistry. Standard deviation is indicated for the four replicates. Significant differences (p-value < 0.05) among groups are indicated by different letters. **(d-f)** Representative overlay of histograms for FL1 intensity of granulocytes (R1) population.

The percentage of myticin C positive granulocytes was also significantly higher in mussels previously bled and then infected with bacteria when compared with mussels only injected with bacteria. Figure 9d-f illustrate representative R1 FL1 profiles for every sample group. Naïve or control mussels injected with FSW showed a granulocyte population positive for myticin C (Figure 9d) that was lost in *Vibrio* challenged mussels without a previous stimulation (Figure 9e). However, the bled/injured mussels after a *Vibrio* challenge did not lose the myticin C labeling (Figure 9f) probably because of the effect of the previous stimulus that prevented the hemocyte population from decaying after a subsequent bacterial infection.

2.4 DISCUSSION

Next generation sequencing technologies are particularly valuable in the study of non-model organisms because they supply the scarce availability of reference genomes, and lead the way to understand many biological processes that could not be studied due to the lack of cellular lines or antibodies. Most of the transcriptomic analysis in the field of bivalves have been conducted using pools of animals or tissues [42,43]. When biological material, such as hemocytes, are scarce, this is a good solution [44–47]. Biological pools help us to eliminate individual differences and allow to focus on the clear patterns of a particular experiment; determining, for example, how bivalves react to a bacterial infection [44–47]. Although this approach is totally correct, we cannot forget that we are dealing with wild animals. They are not laboratory homogeneous strains, and they differ in their genetic and physiological backgrounds. In this work, we were interested in the specific transcriptomic response of individual mussels before and after a bacterial infection.

Unexpectedly, the percentage of DEGs shared between the 3 different mussels after the same stimulus was very small. Each individual had its own repertoire of modulated genes, and this happened in both, controls (response of 3 individuals against FSW) and infected mussels (response of 3 individuals against *V. splendidus*). The high diversity found in mussels for many genes such as antimicrobial

peptides (AMPs), fibrinogen-related proteins (FREPs) or C1q domain containing proteins [19–21] supports the fact that each individual expresses a specific repertoire of transcripts. In this line, an interindividual variability in the basal expression of AMPs was reported in oysters and mussels [13,19,48]. The variability of bivalve responses seems to be related to the antigenic environment of each individual and would also show the genetic diversity of these animals. We wonder if we lose important information when we analyze the immune response in pools of animals instead of using individuals and whether we are underestimating the possibility of the individual immunity as part of a possible social immunity of mussels [49], as their impressive resilience and survival capacities point out.

In addition, the 3 control mussels showed 1,562, 486 and 751 DEGs (using the statistical threshold: FC > |2| and Bonferroni corrected p-value < 0.05), which were modulated by a simple tissue injury (FSW injection). The hemocytes are the bivalve immunocompetent cells, but they are also involved in other physiological processes such as basic homeostasis and wound and shell repair [50,51]. As mussels host diverse and abundant microbial communities, they probably do not respond to all the microorganisms that they are in contact with. A tissue damage could constitute the signal to trigger an immune reaction. Genes related to cell proliferation and migration were overexpressed in control mussels: stathmin, key protein in the cell cycle regulation [52]; low affinity epsilon Fc receptor, with essential roles in defense cell growth and differentiation [53]; cell division cycle-associated protein 2, which controls the cell cycle [54]; signal transducer and transcription activator, member of the STAT family and regulator of cellular immunity, proliferation and differentiation [55]; and structural maintenance of chromosomes protein 4 (SMC4), a protein required to enter into mitotic phase [56]. In contrast, genes related to the regulation of cell death were down-modulated: cell death specification protein 2 (ces-2), with a direct link to the activation of apoptosis [57]; protocadherin Fat 1, a tumor suppressor and inhibitor of cell migration [58]; and GTPase IMAP family member 4 (GIMAP4), member of the GIMAP family which may play a role in defense cells differentiation and apoptosis [59,60]. These results were in agreement with the flow

cytometry results. When an initial injury or bleeding occurred, changes in hemocyte structure were lighter compared with those observed after a bacterial infection. The response of bivalves to tissue injury or different danger signals has not been explored and deserves to be further investigated. These results also highlight the importance of using appropriate controls in our studies. In fact, the acclimation and rest period are important to rule out possible responses due to stress or handling [61]. The usual procedure of injection in the adductor muscle could induce an injury that mussels feel as a danger signal. Therefore, if the aim is to analyze the bivalve immune response against a pathogen, it must be clearly identifying this specific response without the reaction against the tissue damage produced by the injection.

A last finding emerged of the detection of AMPs included in the top 25 most expressed genes from controls (exclusive to controls and not found in infected mussels). To date, the most recognized function of AMPs is the direct killing of microorganisms [62]. In particular, previous works reported biological characteristics of mussel myticins: activity against bacteria [40], activity against molluscan, fish and human viruses [24,41], chemotactic activity [24], etc. Myticins are very diverse, but they are not usually regulated at the transcription level after infection, they used to be stored in hemocyte granules ready to act when needed [41]. However, even with these premises, it was unexpected that myticins were up-regulated in controls. Their expression was increased after a tissue injury, a danger signal. The same response is observed with myticusin in that although it has only been described as an antimicrobial peptide [63], it probably can also be involved in cell proliferation and chemotaxis as myticins. Therefore, the reason why a tissue injury and not a bacterial infection triggers their transcription could be that a bacterial signal (PAMP) would trigger the release of granules to let the AMPs fight the bacteria, and the tissue injury (DAMP) would activate the mechanisms to produce and store these AMPs. This explanation could involve a new and unrecognized function of these AMPs due to exposure to many microorganisms, where wounds matter more than a potential pathogen.

In wild conditions a tissue damage is a possible cause of infection, so it makes sense that these animals could develop mechanisms to

prepare themselves and respond. Interestingly, flow cytometry analysis suggests that myticins could also help in the control of the infection process since the reaction triggered by the bleeding (tissue injury) was enough to avoid the granulocyte and the myticin C+ cell population decrease caused by bacteria, in concordance with the transcriptomic results, which show genes related to cell proliferation. Although further studies will be needed to confirm cell proliferation and hematopoiesis after tissue injury in *M. galloprovincialis*, these results open the door to new research topics regarding innate immunity in bivalves.

2.5 CONCLUSIONS

In summary, we have raised the option of using an individual approach when facing mussel transcriptomics, which may be proper for other organisms as well. We have demonstrated that every single individual expression profile can be very different in terms of the number of genes expressed and in expression magnitude. A careful experimental design should be carried out especially with non-model species, as their individual transcriptomes can be quite diverse. Additionally, some common experimental procedures in bivalves such as shell notching or injection of treatments could trigger unexpected responses such as an unspecific inflammatory response or defense cell proliferation. We found that myticin C may play new roles preparing mussels for future pathogenic processes after a danger signal. A tissue injury is a breach of the first defense barrier, which could be easily followed by a subsequent infection.

More studies should be conducted in the future to understand more about these processes. However, currently, we undoubtedly can say that mussels are anything but simple and that more revelations will appear in the study of "non-model mussel immunity".

2.6 REFERENCES

1. Cole, B.E.; Thompson, J.K.; Cloern, J.E. Measurement of filtration rates by infaunal bivalves in a recirculating flume. Mar Biol. 113, 219–225 (1992). https://doi.org/10.1007/BF00347274.
2. Suttle, C.A. Marine viruses – major players in the global ecosystem. Nat Rev Microbiol. 5, 801–812 (2007). doi:10.1038/nrmicro1750.
3. Stabili, L.; Acquaviva, M.I.; Cavallo, R.A. *Mytilus galloprovincialis* filter feeding on the bacterial community in a Mediterranean coastal area (Northern Ionian Sea, Italy). Water Res. 39, 469–477 (2005). doi:10.1016/j.watres.2004.10.010.
4. Segarra, A.; Pépin J.F.; Arzul, I.; Morga, B.; Faury, N.; Renault, T. Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res. 153, 92–99 (2010). doi:10.1016/j.virusres.2010.07.011.
5. Garcia, C.; Thébault, A.; Dégremont, L.; Arzul, I.; Miossec, L.; Robert, M.; Chollet, B.; François, C.; Joly, J.P.; Ferrand, S.; Kerdudou, N.; Renault, T. Ostreid herpesvirus 1 detection and relationship with *Crassostrea gigas* spat mortality in France between 1998 and 2006. Vet Res. 42, 73 (2011). doi:10.1186/1297-9716-42-73.
6. Garcia, C.; Haond, C.; Chollet, B.; Nerac, M.; Omnes, E.; Joly, J.P.; Dubreuil, C.; Serpin, D.; Langlade, A.; Le Gal, D.; Terre-Terrillon, A.; Courtois, O.; Guichard, B.; Arzul, I. Descriptions of *Mikrocytos veneroïdes* n. sp. and *Mikrocytos donaxi* n. sp. (Ascetosporea: Mikrocytida: Mikrocytiidae), detected during important mortality events of the wedge clam *Donax trunculus* Linnaeus (Veneroida: Donacidae), in France between 2008 and 2011. Parasit Vectors. 11, 119 (2018). doi:10.1186/s13071-018-2692-0.
7. Romero, A.; Costa, M.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. 108, 149–163 (2014). doi:10.3354/dao02701.

8. Domeneghetti, S.; Varotto, L.; Civettini, M.; Rosani, U.; Stauder, M.; Pretto, T.; Pezzati, E.; Arcangeli, G.; Turolla, E.; Pallavicini, A.; Venier, P. Mortality occurrence and pathogen detection in *Crassostrea gigas* and *Mytilus galloprovincialis* close-growing in shallow waters (Goro lagoon, Italy). Fish Shellfish Immunol. 41, 37–44 (2014). doi:10.1016/j.fsi.2014.05.023.
9. Medzhitov, R.; Janeway, C.A.Jr. Innate immunity: the virtues of a nonclonal system of recognition. Cell. 91, 295–298 (1997). doi:10.1016/s0092-8674(00)80412-2.
10. Khalturin, K.; Bosch, T.C. Self/nonself discrimination at the basis of chordate evolution: limits on molecular conservation. Curr Opin Immunol. 19, 4–9 (2007). doi:10.1016/j.coi.2006.11.001.
11. El Chamy, L.; Leclerc, V.; Caldelari, I.; Reichhart, J.M. Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. Nat Immunol. 9, 1165–1170 (2008). doi:10.1038/ni.1643.
12. Berisha, A.; Mukherjee, K.; Vilcinskas, A.; Spengler, B.; Römpp, A. High-resolution mass spectrometry driven discovery of peptidic danger signals in insect immunity. PLoS One. 8, e80406 (2013). doi:10.1371/journal.pone.0080406.
13. Rosani, U.; Varotto, L.; Rossi, A.; Roch, P.; Novoa, B.; Figueras, A.; Pallavicini, A.; Venier, P. Massively parallel amplicon sequencing reveals isotype-specific variability of antimicrobial peptide transcripts in *Mytilus galloprovincialis*. PLoS One. 6, e26680 (2011). doi:10.1371/journal.pone.0026680.
14. Nguyen, T.T.; Hayes, B.J.; Guthridge, K.; Ab Rahim, E.S.; Ingram, B.A. Use of a microsatellite-based pedigree in estimation of heritabilities for economic traits in Australian blue mussel, *Mytilus galloprovincialis*. J Anim Breed Genet. 128, 482–490 (2011). doi:10.1111/j.1439-0388.2011.00948.x.

15. Gerdol, M.; De Moro, G.; Manfrin, C.; Milandri, A.; Riccardi, E.; Beran, A.; Venier, P.; Pallavicini, A. 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. 7, 722 (2014). doi:10.1186/1756-0500-7-722.

16. Moreira, R.; Pereiro, P.; Canchaya, C.; Posada, D.; Figueras, A.; Novoa, B. RNA-Seq in *Mytilus galloprovincialis*: comparative transcriptomics and expression profiles among different tissues. BMC Genomics. 16, 728 (2015). doi:10.1186/s12864-015-1817-5.

17. Murgarella, M.; Puiu, D.; Novoa, B.; Figueras, A.; Posada, D.; Canchaya, C. A First Insight into the Genome of the Filter-Feeder Mussel *Mytilus galloprovincialis*. PLoS One. 11, e0151561 (2016). doi:10.1371/journal.pone.0151561.

18. Moreira, R.; Balseiro, P.; Forn-Cuní, G.; Milan, M.; Bargelloni, L.; Novoa, B.; Figueras, A. Bivalve transcriptomics reveal pathogen sequences and a powerful immune response of the Mediterranean mussel (*Mytilus galloprovincialis*). Mar Biol. 165, 61 (2018). doi:10.1007/s00227-018-3308-0.

19. Costa, M.M.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

20. Romero, A.; Dios, S.; Poisa-Beiro, L.; Costa, M.M.; 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. 35, 334–344 (2011). doi:10.1016/j.dci.2010.10.007.

21. Gerdol, M.; Manfrin, C.; De Moro, G.; Figueras, A.; Novoa, B.; Venier, P.; Pallavicini, A. 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 (2011). doi:10.1016/j.dci.2011.01.018.

22. Pallavicini, A.; Costa M.M.; Gestal, C.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

23. Mitta, G.; Hubert, F.; Noël, T.; Roch, P. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur J Biochem. 265, 71–78 (1999). doi:10.1046/j.1432-1327.1999.00654.x.

24. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS One. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

25. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21, 3674–3676 (2005). doi:10.1093/bioinformatics/bti610.

26. Xia, P.; Xu, X.Y. PI3K/Akt/mTOR signaling pathway in cancer stem cells: from basic research to clinical application. Am J Cancer Res. 5, 1602–1609 (2015).

27. Lampropoulou, V.; Sergushichev, A.; Bambouskova, M.; Nair, S.; Vincent, E.E.; Loginicheva, E.; Cervantes-Barragan, L.; Ma, X.; Huang, S.C.C.; Griss, T.; Weinheimer, C.J.; Khader, S.; Randolph, G.J.; Pearce, E.J.; Jones, R.G.; Diwan, A.; Diamond, M.S.; Artyomov, M.N. Itaconate links inhibition of succinate dehydrogenase with macrophage metabolic remodeling and regulation of inflammation. Cell Metab. 24, 158–66 (2016). doi:10.1016/j.cmet.2016.06.004.

28. Lua, N.H.H.; Medzhitov, R. Food Fight: Role of Itaconate and Other Metabolites in Antimicrobial Defense. Cell Metab. 24, 379–387 (2016). doi:10.1016/j.cmet.2016.08.013.

29. Rico-Bautista, E., Flores-Morales, A.; Fernández-Pérez, L. Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. Cytokine Growth Factor Rev. 17, 431–419 (2006). doi:10.1016/j.cytogfr.2006.09.008.

30. Toubiana, M.; Gerdol, M.; Rosani, U.; Pallavicini, A.; Venier, P.; Roch, P. Toll-like receptors and MyD88 adaptors in Mytilus: complete cds and gene expression levels. Dev Comp Immunol. 40, 158–66 (2013). doi:10.1016/j.dci.2013.02.006.

31. Pauletto, M.; Segarra, A.; Montagnani, C.; Quillien, V.; Faury, N.; Le Grand, J.; Miner, P.; Petton, B.; Labreuche, Y.; Fleury, E.; Fabioux, C.; Bargelloni, L.; Renault, T.; Huvet, A. Long dsRNAs promote an anti-viral response in Pacific oyster hampering ostreid herpesvirus 1 replication. J Exp Biol. 220, 3671–3685 (2017). doi:10.1242/jeb.156299.

32. He, Y.; Jouaux, A.; Ford, S.E.; Lelong, C.; Sourdaine, P.; Mathieu, M.; Guo, X. Transcriptome analysis reveals strong and complex antiviral response in a mollusc. Fish Shellfish Immunol. 46, 131–144 (2015). doi:10.1016/j.fsi.2015.05.023.

33. Rosani, U.; Varotto, L.; Domeneghetti, S.; Arcangeli, G.; Pallavicini, A.; Venier, P. Dual analysis of host and pathogen transcriptomes in ostreid herpesvirus 1-positive *Crassostrea gigas*. Environ Microbiol. 17, 4200–4212 (2015). doi:10.1111/1462-2920.12706.

34. Power, D.; Santoso, N.; Dieringer, M.; Yu, J.; Huang, H.; Simpson, S.; Seth, I.; Miao, H.; Zhu, J. IFI44 suppresses HIV-1 LTR promoter activity and facilitates its latency. Virology. 481, 142–150 (2015). doi:10.1016/j.virol.2015.02.046.

35. Hoffmann, W.; Jagla, W. Cell type specific expression of secretory TFF peptides: colocalization with mucins and synthesis in the brain.

Int Rev Cytol. 213, 147–181 (2002). doi:10.1016/s0074-7696(02)13014-2.

36. Conlon, J.M.; Mechkarska, M.; Kolodziejek, J.; Leprince, J.; Coquet, L.; Jouenne, T.; Vaudry, H.; Nowotny, N.; King, J.D. Host-defense and trefoil factor family peptides in skin secretions of the Mawa clawed frog *Xenopus boumbaensis* (Pipidae). Peptides. 72, 44–49 (2015). doi:10.1016/j.peptides.2015.03.025.

37. Franco-Martínez, L.; Martínez-Subiela, S.; Escribano, D.; Schlosser, S.; Nöbauer, K.; Razzazi-Fazeli, E.; Romero, D.; Cerón, J.J.; Tvarijonaviciute, A. Alterations in haemolymph proteome of *Mytilus galloprovincialis* mussel after an induced injury. Fish Shellfish Immunol. 75, 41–47 (2018). doi:10.1016/j.fsi.2018.01.038.

38. Li, Q.; Wang, X.; Lu, Z.; Zhang, B.; Guan, Z.; Liu, Z.; Zhong, Q.; Gu, L.; Zhou, J.; Zhu, B.; Ji, J.; Deng, D. Polycomb CBX7 directly controls trimethylation of histone H3 at lysine 9 at the p16 locus. PLoS One. 5, e13732 (2010). doi:10.1371/journal.pone.0013732.

39. Fujioka, S.; Niu, J.; Schmidt, C.; Sclabas, G.M.; Peng, B.; Uwagawa, T.; Li, Z.; Evans, D.B.; Abbruzzese, J.L.; Chiao, P.J. NF-kappaB and AP-1 connection: mechanism of NF-kappaB-dependent regulation of AP-1 activity. Mol Cell Biol. 24, 7806–7819 (2004). doi:10.1128/MCB.24.17.7806-7819.2004.

40. Martinez-Lopez, A.; Encinar, J.A.; Medina-Gali, R.M.; Balseiro, P.; Garcia-Valtanen, P.; Figueras, A.; Novoa, B.; Estepa, A. pH-dependent solution structure and activity of a reduced form of the host-defense peptide myticin C (Myt C) from the mussel *Mytilus galloprovincialis*. Mar Drugs. 11, 2328–2346 (2013). doi:10.3390/md11072328.

41. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral activity of myticin C peptide from mussel: an ancient defense against herpesviruses. J Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

42. Zhao, X.; Duan, X.; Wang, Z.; Zhang, W.; Li, Y.; Jin, C.; Xiong, J.; Li, C. Comparative transcriptome analysis of *Sinonovacula constricta* in gills and hepatopancreas in response to *Vibrio parahaemolyticus* infection. Fish Shellfish Immunol. 67, 523–535 (2017). doi:10.1016/j.fsi.2017.06.040.

43. Ren, Y.; Xue, J.; Yang, H.; Pan, B.; Bu, W. Transcriptome analysis of *Ruditapes philippinarum* hepatopancreas provides insights into immune signaling pathways under *Vibrio anguillarum* infection. Fish Shellfish Immunol. 64, 14–23 (2017). doi:10.1016/j.fsi.2017.03.005.

44. 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. 36, 140–149 (2012). doi:10.1016/j.dci.2011.06.012.

45. Tanguy, M.; McKenna, P.; Gauthier-Clerc, S.; Pellerin, J.; Danger, J.M.; Siah, A. Sequence analysis of a normalized cDNA library of *Mytilus edulis* hemocytes exposed to *Vibrio splendidus* LGP32 strain. Results Immunol. 3, 40–50 (2013). doi:10.1016/j.rinim.2013.04.001.

46. 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. 37, 154–165 (2014). doi:10.1016/j.fsi.2014.01.017.

47. Dong, W.; Chen, Y.; Lu, W.; Wu, B.; Qi, P. Transcriptome analysis of *Mytilus coruscus* hemocytes in response to *Vibrio alginnolyficus* infection. Fish Shellfish Immunol. 70, 560–567 (2017). doi:10.1016/j.fsi.2017.08.034.

48. Rosa, R. D.; Alonso, P.; Santini, A.; Vergnes, A.; Bachère, E. High polymorphism in big defensin gene expression reveals presence-absence gene variability (PAV) in the oyster *Crassostrea gigas*.

Dev Comp Immunol. 49, 231–238 (2015). doi:10.1016/j.dci.2014.12.002.

49. Meunier, J. Social immunity and the evolution of group living in insects. Philos Trans R Soc Lond B Biol Sci. 370, 20140102 (2015). doi:10.1098/rstb.2014.0102.

50. Fisher, W.S. Structure and functions of oyster hemocytes in Immunity in Invertebrates (ed. Brehélin, M.) Proceedings in Life Sciences (Springer, 1986).

51. Bachère, E.; Rosa, R.D.; Schmitt, P.; Poirier, A.C.; Merou, N.; Charrière, G.M.; Destoumieux-Garzón, D. The new insights into the oyster antimicrobial defense: Cellular, molecular and genetic view. Fish Shellfish Immunol. 46, 50–64 (2015). doi:10.1016/j.fsi.2015.02.040.

52. Rubin, C.I.; Atweh, G.F. The role of stathmin in the regulation of the cell cycle. J Cell Biochem. 93, 242–250 (2004). doi:10.1002/jcb.20187.

53. Acharya, M.; Borland, G.; Edkins, A.L.; Maclellan, L.M.; Matheson, J.; Ozanne, B.W.; Cushley, W. CD23/FcεRII: molecular multi-tasking. Clin Exp Immunol. 162, 12–23 (2010). doi:10.1111/j.1365-2249.2010.04210.x.

54. Uchida, F.; Uzawa, K.; Kasamatsu, A.; Takatori, H.; Sakamoto, Y.; Ogawara, K.; Shiiba, M.; Bukawa, H.; Tanzawa, H. Overexpression of CDCA2 in human squamous cell carcinoma: correlation with prevention of G1 phase arrest and apoptosis. PLoS One. 8, e56381 (2013). doi:10.1371/journal.pone.0056381.

55. Benekli, M.; Baer, M.R.; Baumann, H.; Wetzler, M. Signal transducer and activator of transcription proteins in leukemia. Blood. 111, 8 (2003). doi:10.1182/blood-2002-04-1204.

56. Kimura, K.; Cuvier, O.; Hirano, T. Chromosome condensation by a human condensin complex in Xenopus egg extracts. J Biol Chem. 276, 5417–5420 (2001). doi:10.1074/jbc.C000873200.

57. Conradt, B.; Wu, Y.C.; Xue, D. Programmed cell death during *Caenorhabditis elegans* development. Genetics. 203, 1533–1562 (2016). doi:10.1534/genetics.115.186247.

58. Hu, X.; Zhai, Y.; Shi, R.; Qian, Y.; Cui, H.; Yang, J.; Bi, Y.; Yan, T.; Yang, J.; Ma, Y.; Zhang, L.; Liu, Y.; Li, G.; Zhang, M.; Cui, Y.; Kong, P.; Cheng, X. FAT1 inhibits cell migration and invasion by affecting cellular mechanical properties in esophageal squamous cell carcinoma. Oncol Rep. 39, 2136–2146 (2018). doi:10.3892/or.2018.6328.

59. Carter, C.; Dion, C.; Schnell, S.; Coadwell, W.J.; Graham, M.; Hepburn, L.; Morgan, G.; Hutchings, A.; Pascall, J.C.; Jacobs, H.; Miller, J.R.; Butcher, G.W. A natural hypomorphic variant of the apoptosis regulator Gimap4/IAN1. J Immunol. 179, 1784–1795 (2007). doi:10.4049/jimmunol.179.3.1784.

60. Heinonen, M.T.; Kanduri, K.; Lähdesmäki, H.J.; Lahesmaa, R.; Henttinen, T.A. Tubulin- and actin-associating GIMAP4 is required for IFN-γ secretion during Th cell differentiation. Immunol Cell Biol. 93, 158–166 (2015). doi:10.1038/icb.2014.86.

61. Thompson, E.L.; Taylor, D.A.; Nair, S.V.; Birch, G.; Coleman, R.; Raftos, D.A. Optimal acclimation periods for oysters in laboratory-based experiments. J Molluscan Stud. 78, 304–307 (2012). doi:10.1093/mollus/eys012.

62. Mitta, G.; Vandenbulcke, F.; Noël, T.; Romestand, B.; Beauvillain, J.C.; Salzet, M.; Roch, P. Differential distribution and defence involvement of antimicrobial peptides in mussel. J Cell Sci. 113, 2759–2769 (2000).

63. Liao, Z.; Huang, W.; Zhou, S.Q.; Wang, X.C.; Liu, H.H.; Fan, M.H.; Wang, R.X.; Gao, P.; Liao, Z. Molecular characterization of a novel antimicrobial peptide from *Mytilus coruscus*. Fish Shellfish Immunol. 34, 610–616 (2013). doi:10.1016/j.fsi.2014.09.019.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 3

IMMUNE TOLERANCE IN *Mytilus galloprovincialis* HEMOCYTES AFTER REPEATED CONTACT WITH *Vibrio splendidus*

Rey-Campos, M.[1]; Moreira, R.[1]; Gerdol, M.[2]; Pallavicini, A.[2,3]; Novoa, B.[1]; Figueras, A.[1]. immune tolerance in *Mytilus galloprovincialis* hemocytes after repeated contact with *Vibrio splendidus*. Front Immunol.10, 1894 (2019). doi:10.3389/fimmu.2019.01894. (Open Access).

[1]Institute of Marine Research (IIM), National Research Council (CSIC), Eduardo Cabello, 6, 36208, Vigo, Spain. [2]Department of Life Sciences, University of Trieste, Trieste, Italy. [3]Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, Trieste, Italy.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 3: Immune tolerance in *Mytilus galloprovincialis* hemocytes after repeated contact with *Vibrio splendidus*

3.1 INTRODUCTION

Although lacking acquired immune response, hemocytes, the immune cells of bivalves, respond to pathogens by chemotaxis, encapsulation, phagocytic activity and releasing oxygen and nitrogen radicals [1], as well as producing effector molecules as the antimicrobial peptides [2,3]. In this sense, transcriptomic information regarding the modulation of the hemocyte immune response in bivalves remains scarce [4–8]. In mussels, several genes related to key immune functions have been characterized over the past few years. These include different molecules involved in specific pathogen recognition, such as C-type lectins [9], C1q domain-containing proteins [10] and proteins with a fibrinogen-related domain [11]. Compared to other bivalves, mussels are also particularly rich in antimicrobial peptides (AMPs), and myticin C is an example of an important immune effector with chemotactic, antibacterial and antiviral activities [12,13]. Lysozyme, which is able to hydrolyze the central components of the bacterial wall, is another key player in arming the mussel immune response [14]. Gerdol and Venier [9] have reviewed the presence and the interplay between the different molecular components of the mussel immune defense system by using information found in public sequence databases.

Because of the lack of antibody production, it has been classically thought that invertebrates could not develop immune memory. But, in the last decade this idea has drastically changed and a new form of

innate immune memory has been defined as "immune priming" or "trained immunity". Because of this, exposure to a non-lethal dose of a pathogen could provide protection against later infections with the same pathogen [1,15]. This enhanced innate immune response after a previous stimulation has been found in many invertebrates and their offspring and it can occur against the same pathogen or even sometimes against a different one [16,17,18]. In molluscs, there are increasing cases of exposures to non-lethal stimuli and subsequent infections that demonstrate the existence of innate immune memory. For example, the gastropod *Biomphalaria glabrata* was protected against a secondary infection of *Schistosoma mansoni*, a natural parasite of these snails [19].

In scallops (*Chlamys farreri*), a priming with *Vibrio anguillarum*, generated protection against a long-term stimulation to the same pathogen [20,21]. This protection was due to the increase of phagocytic and acid phosphatase activity. Furthermore, a shift from a cellular immune response (encapsulation) to a humoral immune response (biomphalysin) was also defined [22]. But the most studied case, as far as bivalves are concerned, is the protection that exerts the poly I:C on the oyster (*Crassostrea gigas*) from subsequent infection with ostreid herpesvirus (OsHV-1) [23–25].

The main objectives of the present study were to characterize the transcriptomic and functional response of mussel hemocytes after an injection with *Vibrio splendidus*, which has been reported to produce some mortalities in mussels [26], and to analyze whether a different type of response could be elicited after a second interaction with the same pathogen. The outcome of this experimental approach might help to reveal the trainability of the mussel immune response and to identify genes associated with this process, scarcely studied in this bivalve.

3.2 MATERIALS AND METHODS

3.2.1 Animals

Adult *M. galloprovincialis* with shells 6–8 cm in length were obtained from a commercial shellfish farm (Vigo, Galicia, Spain) and maintained in open circuit filtered sea water tanks at 15ºC with aeration.

The animals were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, the animals were acclimatized to aquarium conditions for 1 week.

3.2.2 Experimental approach

Twenty mussels were marked and notched in the shell, and hemolymph (500 µl) was withdrawn from the adductor muscle of each mussel with a 0.5mm diameter (25G) disposable needle. The hemolymph sampled at time zero (t0) was centrifuged at 4ºC at 3,000 g for 10 min, and the pellet was resuspended in 500 µl of TRIzol (Invitrogen), immediately homogenized and stored at −80ºC until RNA isolation.

After 1 week, 10 mussels were injected in the adductor muscle with 100 µl of filtered sea water (FSW). The other 10 mussels were injected in the same way with 100 µl of a solution of *V. splendidus* [reference strain, LGP32; 26] at a non-lethal concentration (1×10^7 UFC/mL). One-day post injection (24 hpi), hemolymph (500 µl) was sampled again from individual mussels and centrifuged in the same conditions, and the pellet was resuspended in 500 µl of TRIzol (Invitrogen). Samples were immediately homogenized and kept at−80ºC until RNA isolation.

After 2 weeks, the 10 mussels injected with FSW were injected again with FSW. The mussels previously exposed to *V. splendidus* were injected again with a solution *of V. splendidus* [reference strain, LGP32; 26] at a non-lethal concentration (1×10^8 UFC/mL). One day after the second injection (24 hpi2), hemolymph (500 µl) was sampled again from individual mussels and centrifuged in the same previously described conditions, and the pellet was resuspended in 500 µl of TRIzol (Invitrogen). The samples were immediately homogenized and kept at −80ºC until RNA isolation.

Seven days later (7 d), hemolymph (500 µl) was sampled again from the mussels and centrifuged in the same previously described conditions, and the pellet was resuspended in 500 µl of TRIzol

(Invitrogen). The samples were immediately homogenized and kept at −80ºC until RNA isolation.

3.2.3 *Vibrio splendidus* clearance assessment

The clearance of *V. splendidus* was assessed to make sure that the second injection was made after a complete overcome of a possible infection. cDNA was synthesized from samples taken at t0, 24 hpi, and 7 days after the first infection with 100 ng of total RNA using an NZY First-Strand cDNA Synthesis Kit (nzytech). Gene expression of *V. splendidus* 16S and mussel 18S (used as a reference gene) was analyzed in a Stratagene Mx3005P thermal cycler (Agilent Technologies).

For 16S detection, 5 µl of five-fold-diluted cDNA template was mixed with 0.6 µl of each primer (10µM), 0.4 µl of 16S probe (10µM) and 10 µl of Brilliant III Master Mix 2x Ultrafast (Agilent Technologies) in a final volume of 20 µl. For 18S detection, 1 µl of five-fold-diluted cDNA template was mixed with 0.5 µl of each primer (10µM) and 12.5 µl of Brilliant II SYBR Green (Agilent Technologies) in a final volume of 25 µl. The standard cycling conditions were 95ºC for 10 min, followed by 40 cycles of 95ºC for 15 s and 60ºC for 30 s. All reactions were performed as technical triplicates. The relative expression levels of the genes were normalized using 18S as a reference gene following the Pfaffl method. One-way ANOVA was used to analyze differences in normalized gene expression and bacterial load among the studied samples.

3.2.4 RNA isolation, cDNA production, and Illumina sequencing

RNA isolation was carried out using TRIzol (Invitrogen) according to the manufacturer's protocol. RNA purification was performed after DNase I treatment with the RNeasy Mini Kit (Qiagen). Next, the concentration and purity of the RNA were measured using a NanoDrop ND1000 spectrophotometer, and the RNA integrity was verified with an Agilent 2100 Bioanalyzer (Agilent Technologies). Only the best

RNA samples (in terms of RNA quantity and quality) from four sampling points (t0, 24 hpi, 24 hpi2, 7 d) were chosen for the preparation of cDNA libraries compatible with Illumina sequencing. The chosen samples were from the mussels control 2, control 3, control 4, infected 1, infected 2 and infected 10. A total of 24 samples were selected for sequencing that consisted of 2 conditions, 4 sampling points and 3 biological replicates.

A TruSEq library 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 was then prepared from the fragmented mRNA via reverse transcription, second-strand synthesis and the 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 Qubit system. Paired-end sequencing was performed using an Illumina HiSeqTM 4000 platform by Macrogen Korea.

A schematic representation of the experimental design for the sequenced samples is shown in figure 1. The raw sequencing data have been deposited in the NCBI Short Read Archive database under the accession ID SRP145077.

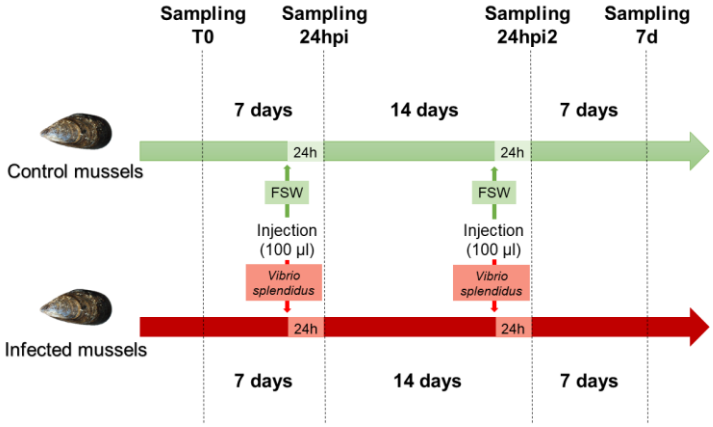

Figure 1. Experimental design used for the stimulation of mussels.

3.2.5 Bioinformatics and RNA-Seq

The CLC Genomics Workbench, v.11.0.1 (CLC Bio; Qiagen), was used to process the raw sequencing output for the *de novo* assembly of the reference transcriptome and to perform the statistical analysis of gene expression by comparing the three biological replicates for the control and infected mussels at different time points. The raw reads were trimmed to remove adaptor sequences, low quality bases (PHRED = 13), and residual sequences shorter than 70 bp. All reads obtained from the 24 libraries were assembled to obtain a complete reference transcriptome with default word size and bubble size parameters. The assembly was cleaned to remove sequences whose origin was mussel ribosomal RNA and mitochondrial mRNAs, as well as contaminant transcripts from *V. splendidus*, ciliates and microalgae. These filtering steps were performed with BLASTn analysis (e-value threshold 1e-10) that were carried out in parallel with the reported assembled mussel genome [27] and the targets mentioned above (the *Vibrio splendidus* genome from strain NCCB 53037, the ciliate *Pseudocohnilembus persalinus* genome, the *Phaeodactylum tricornutum* genome and the *Isochrysis galbana* transcriptome from BioProject PRJNA248394 were used as references). Contigs that produced a more significant hit when compared to the sequences of the putative contaminants than to the mussel genome were discarded. The quality and completeness of the transcriptome were assessed with BUSCO v.3 [28], which was based on the detection of metazoan Benchmarking Universal Single Copy Orthologs (BUSCOs) according to release 9 of OrthoDB.

The reads of each individual mussel and sampling were mapped onto the clean transcriptome with the RNA-Seq tool using the following parameters: mismatch cost = 2, length fraction = 0.8, similarity fraction = 0.8, and maximum hits per read = 10. Differentially expressed genes (DEGs) were identified with a statistical analysis based on generalized linear models and by assuming a negative binomial distribution for the read counts [29]. For each set of comparisons, transcripts with an absolute fold change (FC) value > 2 and an FDR-corrected p-value < 0.05 were considered differentially expressed and retained for further

analysis. To find the DEGs at each time point after the *Vibrio* challenges, the infected samples were compared with the controls. On the other hand, to find primed and tolerized genes, the *Vibrio* challenged samples from the second injection were compared with the challenged samples from the first injection, and the same comparison was made in the controls to confirm that the selected genes were not modulated in control animals.

3.2.6 BLAST annotation, GO assignments, and Enrichment analysis

The transcriptome was functionally annotated with the Blast2GO software [30] by assigning gene ontology (GO) terms based on the significant BLASTx matches found in the UniProt/Swiss-Prot database. To improve the annotation rate, we performed an additional BLASTn analysis against an in-house database, which included all the molluscan sequences present in the NBCI nucleotide database. In both cases, the e-value threshold for annotation was set to 1e-3. Then, functional enrichment analysis of the DEGs (test set) were conducted using the full mussel transcriptome as the reference set. For this purpose, a two-tailed Fisher's exact test was performed with the default parameters and a p-value cut-off of 0.05. The test was performed on the basis of overrepresented biological process (BP) gene ontology terms.

3.2.7 Functional assays: hemocyte distribution, apoptosis, and ROS analysis

The previously described experimental design was repeated using eight biological replicates (each replicate corresponds to one mussel) to determine whether functional immune parameters were also affected by a second exposure to the same bacterial pathogen. Hemolymph was collected from the adductor muscle of the eight individual mussels using a 0.5mm diameter (25G) disposable needle, and the cell concentration was adjusted to 10^6 cells mL^{-1} with FSW.

The hemocyte populations were evaluated by flow cytometry. Two FSC/SSC gates were created, including both the viable granulocyte and hyalinocyte populations. Data were acquired using a FACS Calibur flow cytometer (Becton and Dickinson), and the analysis was carried out using CellQuest software (Becton and Dickinson).

To investigate the effect on cell death rates of receiving a single infection or two subsequent injections of *V. splendidus*, an apoptosis analysis was performed. Hemocytes were centrifuged and resuspended in 1mL of binding buffer (BB1X, Invitrogen). Then, 5 µl of annexin V (Invitrogen) and 5 µl of actinomycin (BD Pharmingen) were added to the cell suspensions. The samples were incubated for 15 min at room temperature in darkness and analyzed by flow cytometry.

The respiratory burst activity of hemocytes was determined by the luminol-enhanced chemiluminescence method (CL) in 96-well-plates. We used 5-amino-2,3-dihydro-1,4-phthalazinedione (luminol, Sigma Aldrich) as a light emitter and phorbol myristate acetate (PMA, Sigma Aldrich) or zymosan A (Sigma Aldrich) to trigger the production of reactive oxygen species (ROS). A stock solution of 0.1M luminol was prepared in dimethyl sulphoxide (DMSO, Sigma Aldrich) and diluted in FSW to obtain the luminol working solution (final concentration of 10 mM). Zymosan A (20 mg mL^{-1}) was diluted in the luminol working solution to obtain a final concentration of 1 mg mL^{-1}. The PMA stock solution (1 mg mL^{-1} in ethanol) was also diluted in the luminol working solution to obtain a final concentration of 1 µg mL^{-1}. One hundred microliters of hemolymph was dispensed into each well of the 96-well plates. After 30 min of incubation at 15ºC, 100 µl of luminol, PMA or zymosan A were added per well. The relative luminescence units (RLU) were measured in a luminometer (Fluoroskan Ascent, Labsystems) six times at intervals of 5 min with an integration time of 1,000 ms for each measurement.

3.3 RESULTS

3.3.1 Assembly and annotation of mussel transcriptome

The sequencing of the individual hemocyte samples yielded an average of 74.11 million raw reads for each of the 24 libraries. The trimming procedure removed, on average, 0.83% of the raw reads, and a total of 1,778 million reads were assembled into a reference mussel transcriptome containing 260,664 contigs with an average length of 512 bp. The reference transcriptome was highly complete, as just 2% of metazoan BUSCOs were absent and displayed a fragmentation rate equal to 22%, what was expected for such a highly heterozygous species (Table 1).

Two different BLAST approaches were used to annotate the assembled transcriptome. In brief, 42.64% of the contigs were found to have a significant match in the custom database, which included all the mollusc nucleotide sequences available from NCBI, and Blast2GO was used to annotate 19.93% of the contigs through a BLASTx search of UniProt/Swiss-Prot. Based on these results, gene ontology (GO) terms were assigned to 23.35% of the contigs. Table 1 shows the sequencing output of all the samples and the main metrics of the transcriptome assembly and annotation.

Table 1. Summary of the transcriptome bioinformatics pipeline.

Sample	Raw reads	Trimmed reads
C2 t0	78,426,948	99.59%
C3 t0	44,346,854	98.03%
C4 t0	100,814,198	99.59%
I1 t0	17,696,894	98.08%
I2 t0	93,114,098	99.60%
I10 t0	96,780,602	99.64%
C2 24h	95,296,484	99.49%
C3 24h	51,708,988	97.62%
C4 24h	92,661,282	99.65%
I1 24h	52,102,302	98.77%
I2 24h	90,965,875	99.52%
I10 24h	99,262,970	99.55%
C2 24h2	93,665,372	99.96%
C3 24h2	47,632,114	99.09%
C4 24h2	83,629,996	99.62%
I1 24h2	25,833,124	98.25%
I2 24h2	90,636,926	99.60%
I10 24h2	101,411,946	99.64%
C2 7d	78,528,396	99.55%
C3 7d	46,511,774	99.36%
C4 7d	95,180,866	99.69%
I1 7d	13,593,322	97.43%
I2 7d	90,614,296	99.31%
I10 7d	98,236,250	99.57%
Assembly statistics		
Contigs	260,664	
Range contig length	200-15,624	
Average contig length	512	
N50	576	
Complete metazoan BUSCOs	740/978 (75.69%)	
Fragmented metazoan BUSCOs	218/978 (22.29%)	
Missing metazoan BUSCOs	20/978 (2.04%)	
Blast		
Contigs with hit in UniProt/Swiss-Prot	19.93%	
Contigs with hit in molluscs database	42.64%	
GO analysis		
Annotated contigs	23.35%	
KEGG analysis		
Pathway assigned contigs	7.27%	

3.3.2 Transcriptomic response after injection and reinjection with the same pathogen

A differential gene expression analysis was carried out to gain insights into the dynamics of the transcriptional response of mussel hemocytes to an experimental infection with *V. splendidus*. First of all, the clearance of *V. splendidus* by the injected mussels was confirmed 7 days after the first injection: *V. splendidus* detection increased 24 hpi and was rapidly controlled 7 days after the injection, returning to control levels (Figure 2).

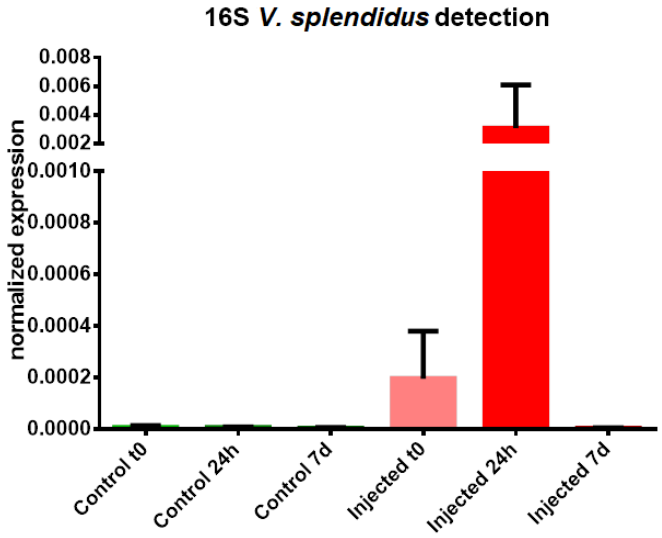

Figure 2. 16S *V. splendidus* detection by qPCR. The relative expression levels of the genes were normalized using mussel 18S as a reference gene.

To analyze the transcriptomic response at each sampling point, the *Vibrio* injected animals were compared to control animals (FSW-injected) (Figure 3a). The monitoring of the transcriptional profiles enabled us to assess whether a second interaction with the same pathogen could elicit a different type of response compared to the response elicited by the first injection. A total of 1,216 differentially expressed genes (DEGs) were detected 24 h after the first injection (24 hpi). However, this number dramatically decreased to 236 DEGs 24 h after the second injection (24 hpi2) and dropped further to 80 DEGs

when the transcriptional profiles were compared 7 days after reinjection (7 d) (Figure 3b).

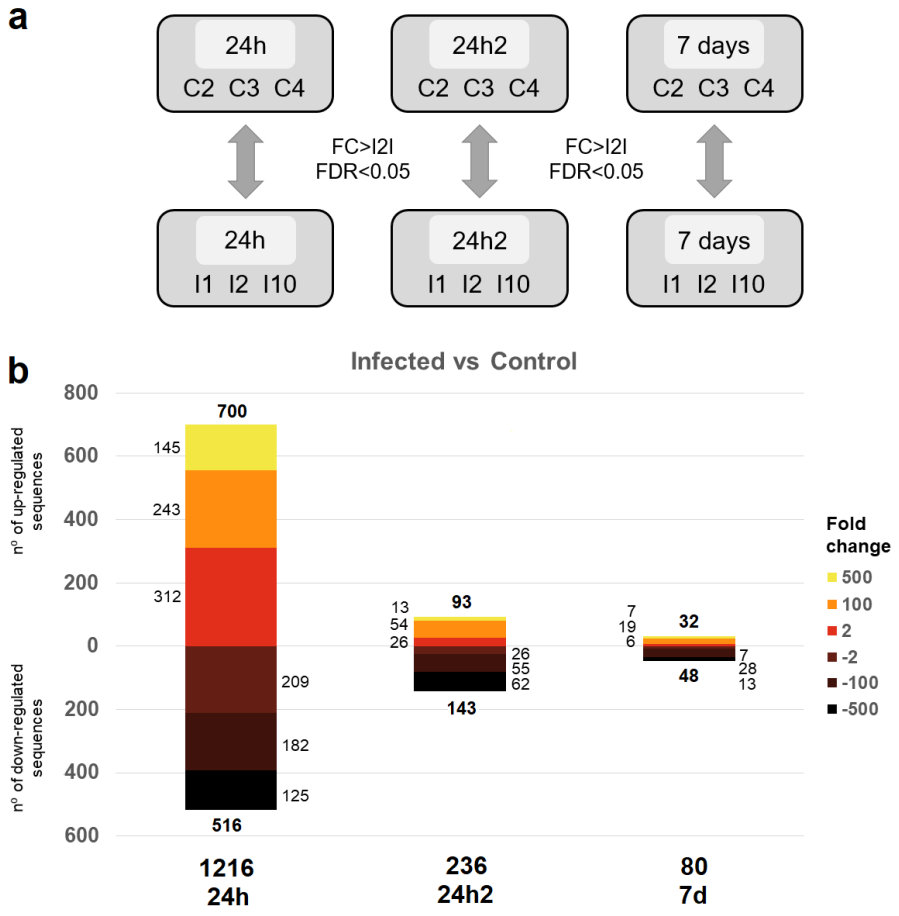

Figure 3. (a) Representative scheme of the comparisons made for the differential expression analysis between injected and control individuals at the analyzed sampling points. The thresholds used to detect statistical significance were a fold change (FC) > |2| and a FDR < 0.05. **(b)** Distribution of the response magnitudes of the DEGs. Statistically significant gene modulation is shown according to intensity (fold-change) and sense (up- and down-regulation).

To detect any significant alteration in biological pathways, a Fisher's exact test was performed. An enrichment analysis of the GO

annotations associated with the DEGs at each sampling point (24 hpi, 24 hpi2, and 7 d) was conducted. The 30 most significantly enriched GO terms for each sampling point are shown in figure 4. After the first injection (24 hpi), GO terms related to the immune system were found, such as those related to the regulation of innate immune response, inflammatory response, cell migration and defense response to bacteria. After the second injection (24 hpi2), genes related to the inflammatory response seemed to be also modulated and were represented in processes involved in the regulation of the NF-kB signaling pathway. Moreover, processes involved in defense response to bacteria and fungi, negative regulation of ROS, apoptosis and glucose homeostasis appeared to be regulated after reinjection. The last sampling point (7 d) showed a modulation of genes related to GO terms involved in neural processes (long-term memory and learning), tissue regeneration (cell population proliferation and proteoglycan, glycosaminoglycan, mucopolysaccharide, and collagen metabolism) and the resolution of infections (oxidation-reduction processes and defense response to pathogens).

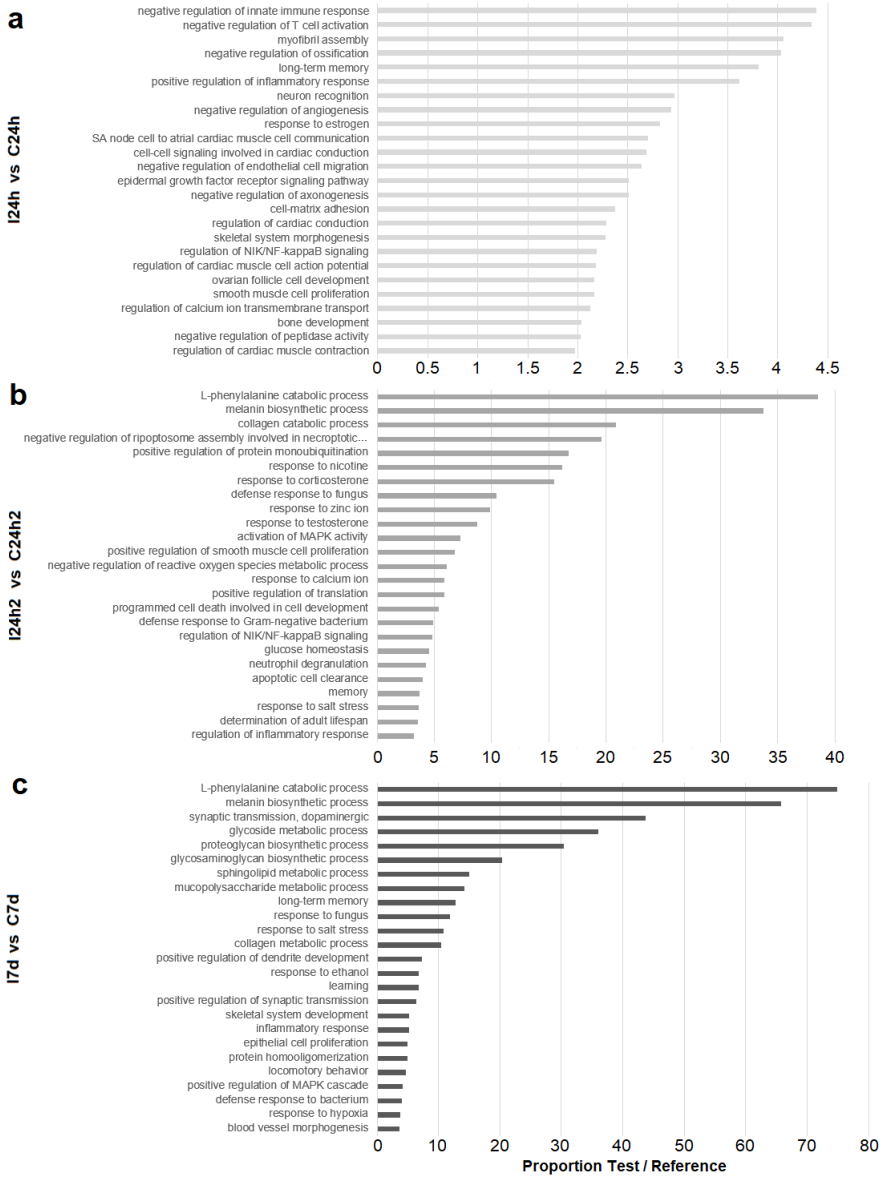

Figure 4. Enrichment analysis of DEGs. Bars represent the proportions between the percentages of sequences in the test set (DEGs list) and the reference set (global transcriptome). **(a)** Biological processes (BP) overrepresented in infected mussels 24 h after the first infection. **(b)** BP overrepresented 24 h after the second infection. **(c)** BP overrepresented 7 days after the second infection.

The most highly expressed genes at each sampling point are shown in Table 2 (complete information is available in https://www.frontiersin.org/articles/10.3389/fimmu.2019.01894/full#supplementary-material). After the first exposure, several genes showed high expression values that significantly decreased as the experiment progressed (reinjection and 7 d). This was the case for perlucin-like protein, which is directly involved in pathogen recognition, the spore cortex-lytic enzyme, which can destroy the bacterial cell wall, and the henna protein, which is important for melanization; all of these genes could play crucial roles in the killing and sequestration of invading pathogens. These genes reached very high fold-change values after the first injection, they decreased after reinjection and exhibited their lowest values 7 days after reinjection. Other interesting genes related to recognition (lectin, neurocan, and galaxin), acute phase response (HSP70 and sacsin), antimicrobial response (apextrin) and apoptosis (caspase 3 and the GTPase IMAP family member 4) were up- or down-regulated in a balanced manner (15 up- and 10 down-regulated). However, one day after the second challenge, the majority of the most highly regulated genes were down-regulated (23 down- and 2 up-regulated). For example, antimicrobial peptides such as defensin MGD-1 or myticin B were not differently regulated after the first injection, but they were indeed inhibited after reinjection (FC −1,360 and -2107 respectively) and 7 days after the second injection (Myticin B FC −3,745).

Table 2. Top 20 DEGs at each sampling point. FC, fold change.

I24h vs C24h		I24h2 vs C24h2		I7d vs C7d	
FC	**Description**	**FC**	**Description**	**FC**	**Description**
-6068,31	Caspase-3	-20257,80	Vitellogenin	-3745,45	Myticin B
-5928,63	CHRNA2	-7278,01	TCAF2	-1031,26	Apextrin-like protein 1
-5633,81	TIMP3	-5276,43	Latrophilin Cirl	-956,71	Neurocan core protein
-5114,18	GIMAP4	-5076,28	COX3	-712,63	L-cys desulfhydrase 1
-3240,46	Hepatic lectin	-4899,42	PEPCK	-693,30	GRHPR
-3109,94	Apextrin-like protein 1	-4849,28	Bacterial hemoglobin	-620,87	C1QL3
-2965,72	Nephrin	-3940,14	BHMT	-607,11	TIMP3
-2778,00	Neurocan core protein	-3754,37	SCARB2	-596,49	SLC25A3
-2611,24	Sacsin	-3724,18	Multi-CRP-I 3	-562,79	GRHPR
2725,96	ALO1	-3590,02	40S ribosomal protein SA	-510,31	CHRNA2
3084,32	OAS1	-3381,63	Venom allergen 5	-407,37	Cys protease RD21B
3102,00	WFDC2	-3007,79	Vitellogenin-2	-246,99	Multi-CRP-I 3
3139,69	Shell protein-5	-2451,55	TIMP3	-203,99	Hepatic lectin
3251,22	Galaxin	-2405,10	Apextrin-like protein 1	336,24	Perlucin-like protein
3255,70	Netrin receptor DCC	-2107,39	Myticin B	359,33	Peroxidasin homolog
3294,33	Nephrin	-1566,39	Papilin	419,55	Spore cortex-lytic enzyme
4011,09	Spore cortex-lytic enzyme	-1482,80	MDH1	427,42	Phenylalanine-4-hydroxylase
4027,72	Protein henna	-1466,81	Heme-binding protein 2	735,42	Protein henna
4927,82	HSPA12B	-1360,83	Defensin MGD-1	915,43	Collagen alpha-2 (VIII) chain
5793,60	LRP6	1451,53	Spore cortex-lytic enzyme	4103,11	Lysozyme

3.3.3 Variations in the response after the second encounter with the same pathogen

Next, we looked for trainable genes with different expression values after reinjection compared to those after the first injection. We compared the transcriptomes of the challenged animals after the first and second exposures and selected primed genes, which were those that increased the expression values after the second encounter with the same pathogen, or tolerized genes, which were those that showed decreased expression after reinjection (Figure 5a).

A schematic representation of the expression behavior of these genes during the first and second injections with regard to naïve animals is shown in figure 5b. Thirty-nine genes showed significantly increased expression after previous stimulation with *V. splendidus*, and 31 showed decreased expression. The expression levels of all these genes did not change in control animals after the first or second stimulation (Figure 5c).

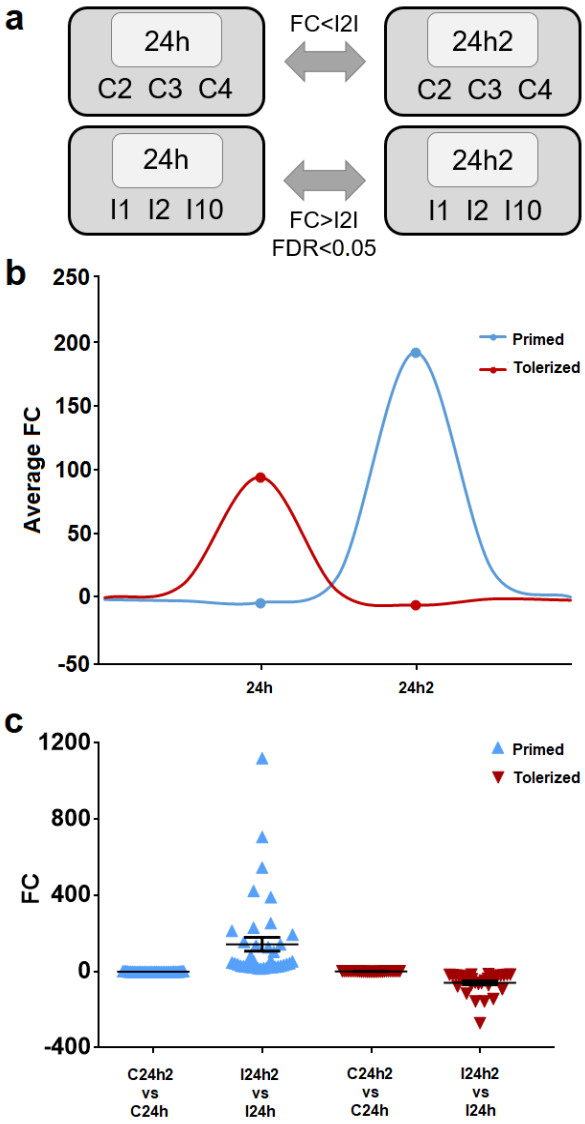

Figure 5. (a) Representative scheme of the comparisons made to select the primed and tolerized genes. The statistical parameters used as thresholds were an FDR-corrected p-value < 0.05 and a fold change (FC) > |2| in the case of *Vibrio* injected individuals and a FC < |2| in control mussels. **(b)** Mean of the tendency of the primed and tolerized genes. **(c)** Fold change of each primed and tolerized gene and their respective controls. The mean and standard error of all represented genes is also shown.

The list of the annotated primed and tolerized DEGs is shown in Table 3. The regulation of these genes suggests an attenuation of inflammation, a decrease in radical oxygen species (ROS) production and the inhibition of apoptosis in the second contact with *V. splendidus*.

Table 3. Identified primed and tolerized DEGs.

		Primed DEGs 24h2 vs 24h
Contig	**FC**	**Description**
Mg_5596	423,00	Myomodulin neuropeptides
Mg_14406	230,36	Mitochondrial uncoupling protein 2-like: UCP2
Mg_1335	193,45	Tetraspanin-7
Mg_9808	155,40	Mitochondrial uncoupling protein 2-like: UCP2
Mg_4328	142,49	Furin-like protease: KPC-1
Mg_30535	136,05	Tubulin beta chain
Mg_14134	128,00	Plasminogen
Mg_3665	57,18	Cytochrome P450 3A24
Mg_50120	34,93	DNA repair protein complementing XP-A cells: XPA
Mg_24689	30,88	SLC12A8
Mg_16026	30,15	Ropporin-1-like protein
Mg_21973	27,24	Inhibitor of p53-induced apoptosis-beta
Mg_13532	26,93	Zinc finger protein Eos
Mg_3349	25,19	Oxidative stress-induced growth inhibitor 1: OKL38
Mg_13357	23,18	Apolipoprotein(a)
Mg_41851	22,52	Tetratricopeptide repeat protein 38
Mg_36957	22,50	Hydrocephalus-inducing protein
Mg_32198	22,28	SLC12A8
Mg_45429	18,20	Alpha-L-fucosidase
Mg_15455	17,40	Solute carrier family 46 member 3
Mg_13725	14,90	NAD kinase
		Tolerized DEGs 24h2 vs 24h
Contig	**FC**	**Description**
Mg_49351	-157,33	Usherin
Mg_53728	-69,50	NFX1-type zinc finger-containing protein 1
Mg_88014	-40,66	Regulator of nonsense transcripts 1: UPF1
Mg_50460	-22,38	NFX1-type zinc finger-containing protein 1
Mg_38773	-21,16	Nephrin
Mg_1952	-19,36	H/ACA ribonucleoprotein complex subunit 4: DISKERIN
Mg_2166	-18,87	FARSB
Mg_39783	-17,22	Nuclear migration protein: nudC
Mg_7966	-16,66	PheS
Mg_13367	-16,20	Dual serine/threonine and tyrosine protein kinase: RIP5
Mg_14736	-13,65	FARSB

Primed genes, with increased expression in the second encounter, such as those encoding myomodulin neuropeptides, furin-like protease (KPC-1), and plasminogen/apolipoprotein(a), were directly or secondarily related to the control and inhibition of inflammatory processes. Moreover, there was high expression of genes involved in the inhibition of ROS, such as mitochondrial uncoupling protein 2-like (UCP2), oxidative stress-induced growth inhibitor 1 (OKL38), and NAD kinase. Finally, genes involved in the reduction of cell death (inhibitor of p53- induced apoptosis-beta) and DNA repair (DNA repair protein complementing XP-A cells) were also present in our set of primed genes. However, tolerized genes, with a reduced expression level after the second encounter, were associated with the activation of apoptosis and inflammatory response and included regulator of nonsense transcripts 1 (UPF1), nephrin, H/ACA ribonucleoprotein complex subunit 4 (DISKERIN), nuclear migration protein (nudC), and the dual serine/threonine and tyrosine protein kinase (RIP5).

3.3.4 Priming induces the modification of functional hemocyte response

Flow cytometry was used to better understand how two consecutive injections of *V. splendidus* affected mussel hemocytes. Two cell populations, granulocytes (R2) and hyalinocytes (R3), were well-defined in the control group (mussels injected with FSW at both sampling points) (Figure 6a). When mussels were injected with *V. splendidus*, the cell population structure was altered, and it was almost impossible to establish two separate populations of granulocytes and hyalinocytes. However, if mussels were stimulated with *V. splendidus* and received a second injection with the same pathogen, the hemocyte population structure was restored to that found in naïve mussels and showed a distribution similar to that in the controls. Quantitatively, the numbers of granulocytes and hyalinocytes were significantly reduced in mussels injected once with the bacteria compared to control animals. However, when mussels had been previously injected and received a second bacterial challenge, both cell types reached similar values to those found in the control (Figure 6b).

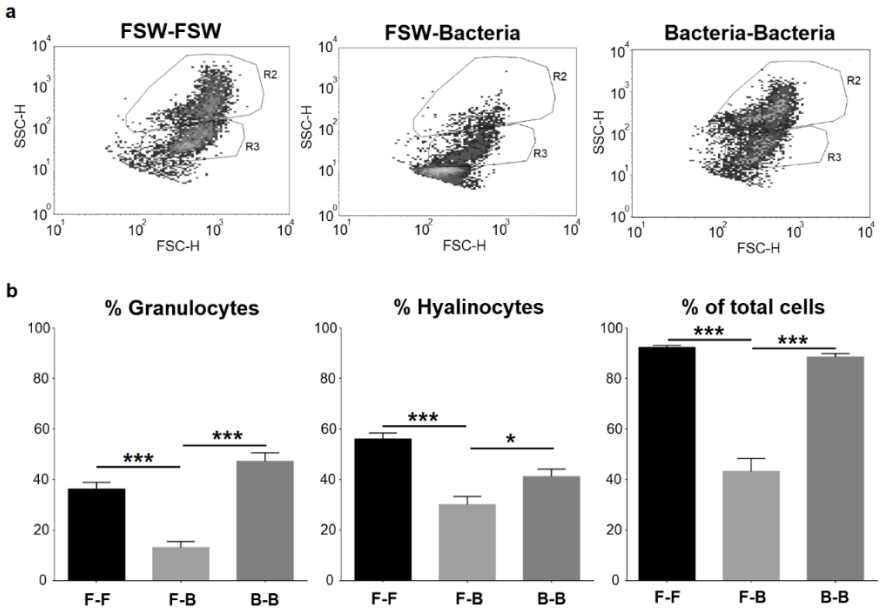

Figure 6. FACS analysis of the cell population distribution of mussel hemocytes after subsequent FSW/bacteria injection. **(a)** Dot plot of a total count of 100,000 events for hemolymph. R2, granulocytes; R3, hyalinocytes. **(b)** Statistical analysis of the hemocyte population in the different experimental conditions. Asterisks show the statistical significance: *p < 0.05 and ***p < 0.0001. Subsequent stimulations are indicated as follows: F-F, FSW and FSW; F-B, FSW and bacteria; B-B, bacteria and bacteria.

To confirm the results of the transcriptomic analysis that suggested the inhibition of respiratory burst activity after the second injection, we looked closely at the expression values (TPM values) of some representative genes (OKL38 and UCP-2) in individual mussels. All challenged animals exhibited a significant increase in the expression of these two antioxidant genes after reinjection (24 hpi2) (Figure 7a). We also analyzed ROS production in hemocytes from treated mussels. Respiratory burst activity when there was no triggering molecule or that was triggered by PMA was notably decreased after the second injection (Figure 7b), supporting the results of the transcriptomic analysis. Respiratory burst activity mediated by zymosan A did not show significant differences among the three groups of mussels (FSW-FSW,

FSW-bacteria, and bacteria-bacteria), which was probably due to the strong stimulating effect of zymosan A that masked the natural response [31].

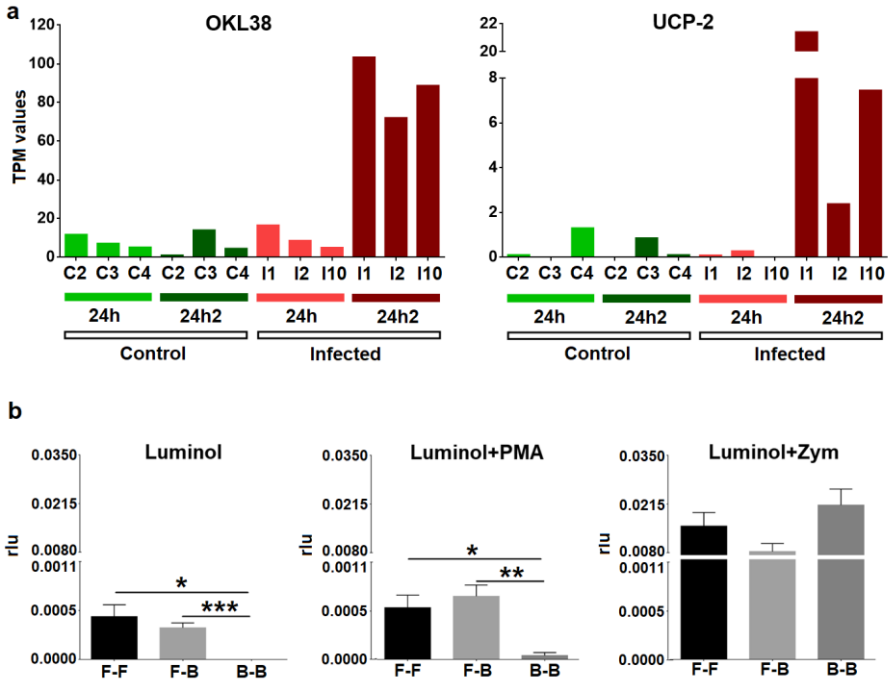

Figure 7. (a) Expression values (TPM) for two representative inhibitors of oxidative stress (OKL38 and UCP-2) in the six studied individuals (C2, C3, C4, I1, I2, and I10) at two experimental time points (24 hpi and 24 hpi2). **(b)** Respiratory burst assay. Bars represent the mean percentage and standard error of the relative luminescence units (rlu) for the total hemocytes in eight individual mussels. Asterisks show the statistical significance: *p < 0.05; **p < 0.001; and ***p < 0.0001. Subsequent stimulations are indicated as follows: F-F, FSW and FSW; F-B, FSW and bacteria; B-B, bacteria and bacteria.

Apoptosis was another central process regulated in the transcriptomic analysis. This process appeared strongly inhibited after the second *Vibrio* injection (24 hpi2); therefore, we analyzed the expression of cell death inducers such as UPF1, RIP5, and nephrin in individual mussels (Figure 8a) and confirmed the results observed in

the global analysis. We next performed an experiment to analyze whether two subsequent challenges resulted in changes in cell death. The number of apoptotic granulocytes and hyalinocytes was significantly increased when mussels were injected with *V. splendidus*. However, if mussels were injected two times with a resting phase in between, the number of apoptotic cells was similar to that detected after the first injection without any further increase in the apoptotic rate (Figure 8b).

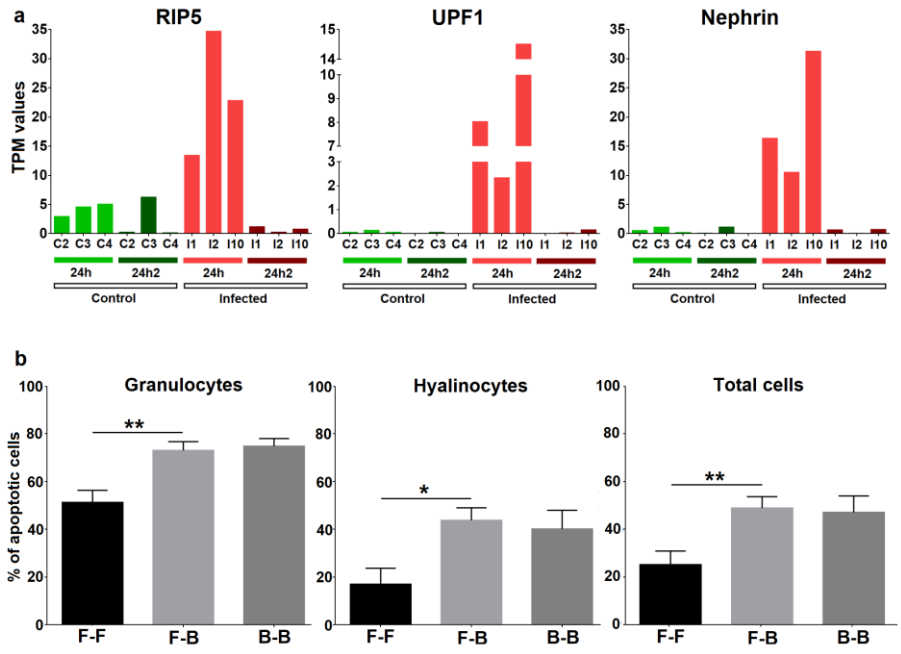

Figure 8. (a) Expression values (TPM) for three representative apoptosis inducers (RIP5, UPF1, and nephrin) in the 6 studied individuals (C2, C3, C4, I1, I2, and I10) at two experimental time points (24 hpi and 24 hpi2). **(b)** Apoptosis analysis. Bars represent the mean percentage and standard error of apoptotic cells in total hemocytes, hyalinocytes and granulocytes in eight individual mussels. Asterisks show the statistical significance: *p < 0.05 and **p < 0.001. Subsequent stimulations are indicated as follows: F-F, FSW and FSW; F-B, FSW and bacteria; B-B, bacteria and bacteria.

Our transcriptomic and functional results suggest that there is a modulation of the immune response after a second encounter with the same pathogen. Primed genes are involved in the resolution of the inflammatory process and the inhibition of ROS; however, repressed transcripts are related to inflammatory reactions and oxidative stress. There is a shift toward an anti-inflammatory response that attempts to minimize the damage caused by the second encounter with *V. splendidus* (Figure 9).

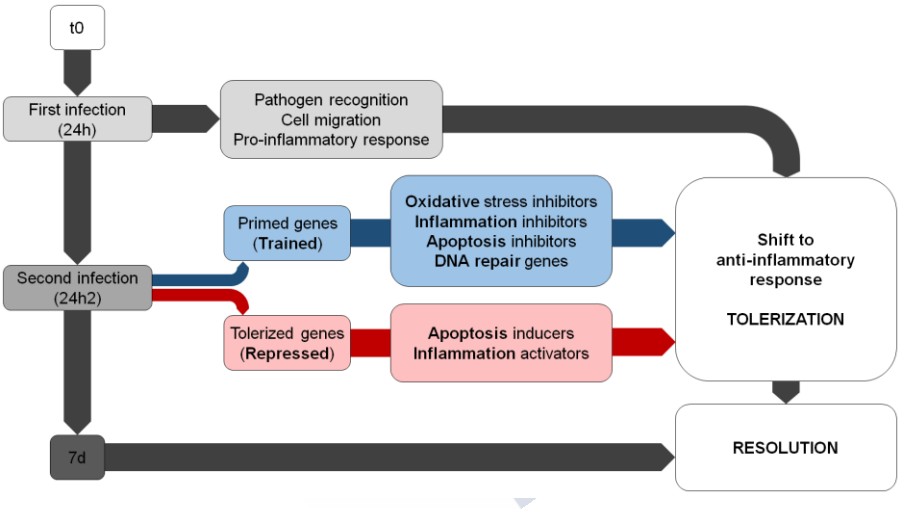

Figure 9. Summary of the main biological processes regulated during the subsequent *Vibrio* challenges. Note that the trained genes shift in terms of response from pro- to anti-inflammatory.

3.4 DISCUSSION

One of the characteristics of the innate immune system is the lack of immunological memory. However, in recent years, there is increasing evidence that innate immune cells can become reprogrammed to develop immunological memory after previous encounters with non-self-molecules [32–34]. Bivalves, like all invertebrates, do not have an adaptive immune system and, due to their status as filter-feeding animals, are constantly in contact with microorganisms and

environmental pollutants. Our results suggest that mussels may control the magnitude of their immune response, which allows them to deal with the continuous exposure to potential pathogens. A continuous reaction in these animals against all pathogenic and non-pathogenic microbes would potentially result in a constant state of inflammation that may be detrimental for the organism.

After the first *V. splendidus* exposure, there was an initial response represented by a high number of modulated genes, which decreased after the second exposure and almost returned to a basal level at the end of the experiment. The decrease in the number of differentially expressed genes after the second injection is in concordance with previous findings in *Crassostrea gigas* [35], in which specific protection against a viral infection was achieved after poly I:C priming and a later exposure to the pathogen did not trigger an antiviral response. This suggests that in our experimental design, the primary challenge with a pathogen triggered immune processes that could be reprogrammed afterwards.

When comparing the processes significantly enriched during the two subsequent encounters, a shift from an inflammatory to an anti-inflammatory state can be observed. After the first injection, the expected response would involve the positive regulation of the inflammatory response and the migration of hemocytes, as seen in other bivalves [1]. However, a strong decrease in the number of DEGs and the strict control of inflammation, which is a potentially harmful process, could be observed after the second challenge, which was possibly a consequence of the inhibition of the NF-kB signaling pathway [36]. Taking into account that *Vibrio splendidus* was used at a sublethal dose and also that the infection was cleared before the second injection, it is unlikely that bacteria virulence factors have been responsible for the regulation of the inflammatory response. Also, we cannot know if the response after the second injection would be the same if a different pathogen was injected. These aspects should be further investigated.

Inflammation is critical in the response against infection; however, this process cannot last for a long time, and a return to a non-

inflammatory state requires specific suppressor molecules [37]. The expression of primed and tolerized genes in the mussel transcriptome suggests a combined response that attempts to control and limit three key immune processes: inflammation, ROS production, and apoptosis. The primed genes are involved in the attenuation of inflammation by regulating the transport of ions (myomodulin neuropeptides; [38,39]), inhibiting NF-kB via ubiquitination (KPC-1; [40]), or inhibiting inflammatory pathways (plasminogen/apolipoprotein(a); [41–43]). On the other hand, the tolerized genes (inhibited in the second challenge) are involved in the regulation of cell death (RIP5; [44]).

As in the case of the inflammatory process, an excess of reactive oxygen species can be harmful to the organism. At physiological levels, ROS are involved in intracellular signaling and defense, but uncontrolled production yields oxidative stress. Therefore, ROS are strictly controlled in all organisms to prevent self-inflicted damage [45,46]. The production of ROS is a well-characterized defense process in bivalves [47], but after the second exposure to the same pathogen, mussels seemed to actively control oxidative stress by inhibiting respiratory burst activity with the expression of genes such as UCP2 [48], OKL38 [49], and NAD kinase [50]. It seems that the control of oxidative stress is one of the central modulated processes after repeated encounters. Accordingly, at the functional level, we observed that the ROS levels were reduced after two subsequent *V. splendidus* challenges, suggesting that hemocytes could prevent an uncontrolled respiratory burst. Moreover, hemocytes seem to avoid cellular impairment caused by DNA damage resulting from previous cited oxidative stress [51], with the overexpression after the second injection of the DNA repair protein XPA [52].

A strong link between oxidative response and apoptosis has been reported [53,54]. The tolerized genes also include several modulators of apoptosis. These genes, which have a predominantly pro-apoptotic function, are inhibited after the second exposure and include UPF1 [55], nephrin [56], DISKERIN [57], RIP5 [44], and nudC [58]. At the functional level, the number of apoptotic cells was significantly increased after a single injection and was maintained, rather than increased, after the second injection.

In addition, the loss of the hemocyte population distribution after the first challenge and the restoration to normal conditions after the second challenge are in concordance with the presence of a priming process and reflect hemocyte recovery after the previous *V. splendidus* encounter. It seems obvious that it exists a change in hemocyte populations, as it is in concordance with previous priming results in molluscs [59–62], and it should be further explored.

Eventually, 7 days after the second exposure, the resolution of the infection occurred. Processes related to tissue regeneration and learning are represented in our results, including the up-regulation of proliferation (C3a anaphylatoxin chemotactic receptor, skin secretory protein xP2, sushi, nidogen, and EGF-like domain-containing proteins), maintenance of the extracellular matrix (collagen, techylectin, perlucin, neurocan, and glycoproteins), and learning (phenylanaline-4-hydroxylase/henna protein). The presence of neural processes related to the generation of memory at all sampling points is remarkable. Although this might be due to the strong bias in the GO database in terms of model organisms, we cannot discard evidence of a process that occurs during infection. The protein related to these GO terms, which is involved in the melanization cascade in invertebrates, could also be related to ancient cognition and behavior mechanisms, which are known to occur in invertebrates [63]. In any case, it seems that certain processes, such as the metabolism of phenylalanine, could be involved in the generation of innate immune memory. Previous studies in mammals have shown the impairment of cognitive function due to phenylalanine hydroxylase deficiency [64–66], showing a possible link to learning processes and the evolutionary conservation of this mechanism.

3.5 CONCLUSIONS

In summary, the immune responses of *M. galloprovincialis* after the first and second encounter with *V. splendidus* were different. The analysis of the differentially expressed genes suggests that, after the second contact with the bacteria, mussel hemocytes attempted to control and resolve the inflammatory response to avoid subsequent

DNA damage and cell death. There appears a tightly regulated response shifting from a pro-inflammatory response to an anti-inflammatory and probably regenerative phenotype. In conclusion, these results indicate the existence of a secondary immune response in mussels oriented to tolerate infection by inducing anti-inflammatory processes to minimize tissue damage.

3.6 REFERENCES

1. Allam, B.; Raftos, D. Immune responses to infectious diseases in bivalves. J Invertebr Pathol. 131, 121–36 (2015). doi:10.1016/j.jip.2015.05.005.

2. Costa, M.M.; Prado-Alvarez, M.; Gestal, C.; Li, H.; Roch, P.; 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. 26, 515–23 (2009). doi:10.1016/j.fsi.2009.02.001.

3. Venier, P.; Varotto, L.; Rosani, U.; Millino, C.; Celegato, B.; Bernante, F.; Lanfranchi, G.; Novoa, B.; Roch, P.; Figueras, A.; Pallavicini, A. Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics. 12, 69 (2011). doi:10.1186/1471-21 64-12-69.

4. Renault, T.; Faury, N.; Barbosa-Solomieu, V.; Moreau, K. Suppression substractive hybridisation (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster, *Crassostrea gigas*, challenged with Ostreid herpesvirus 1. Dev Comp Immunol. 35, 725–35 (2011). doi:10.1016/j.dci.2011.02.004.

5. Li, J.; Zhang, Y.; Zhang, Y.; Liu, Y.; Xiang, Z.; Qu, F.; Yu, Z. Cloning and characterization of three suppressors of cytokine signaling (SOCS) genes from the Pacific oyster, *Crassostrea gigas*. Fish Shellfish Immunol. 44, 525–32 (2015). doi:10.1016/j.fsi.2015.03.022.

6. Wang, Z.; Wang, B.; Chen, G.; Jian, J.; Lu, Y.; Xu, Y.; Wu, Z. Transcriptome analysis of the pearl oyster (*Pinctada fucata*) hemocytes in response to *Vibrio alginolyticus* infection. Gene. 575, 421–8 (2016). doi:10.1016/j.gene.2015.09.014.

7. Wei, J.; Liu, B.; Fan, S.; Li, H.; Chen, M.; Zhang, B.; Su, J.; Meng, Z.; Yu, D. Differentially expressed immune-related genes in hemocytes of the pearl oyster *Pinctada fucata* against allograft identified by transcriptome analysis. Fish Shellfish Immunol. 62, 247–56 (2017). doi:10.1016/j.fsi.2017.01.025.

8. Moreira, R.; Balseiro, P.; Forn-Cuní, G.; Milan, M.; Bargelloni, L.; Novoa, B.; Figueras, A. Bivalve transcriptomics reveal pathogen sequences and a powerful immune response of the Mediterranean mussel (*Mytilus galloprovincialis*). Mar Biol. 165, 61 (2018). doi:10.1007/s00227-018-3308-0.

9. Gerdol, M.; Venier, P. An updated molecular basis for mussel immunity. Fish Shellfish Immunol. 46, 17–38 (2015). doi:10.1016/j.fsi.2015.02.013.

10. Gerdol, M.; Manfrin, C.; De Moro, G.; Figueras, A.; Novoa, B.; Venier, P.; Pallavicini, A. 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 (2011). doi:10.1016/j.dci.2011.01.018.

11. Romero, A.; Dios, S.; Poisa-Beiro, L.; Costa, M.M.; 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. 35, 334–44 (2011). doi:10.1016/j.dci.2010.10.007.

12. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS ONE. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

13. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral activity of myticin C peptide from mussel: an ancient defense against herpesviruses. J Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

14. Li, H.; Parisi, M.G.; Toubiana, M.; Cammarata, M.; Roch, P. Lysozyme gene expression and hemocyte behaviour in the Mediterranean mussel, *Mytilus galloprovincialis*, after injection of various bacteria or temperature stresses. Fish Shellfish Immunol. 25, 143–52 (2008). doi:10.1016/j.fsi.2008.04.001.

15. Milutinovic, B.; Kurtz, J. Immune memory in invertebrates. Semin Immunol. 28, 328-42 (2016). doi:10.1016/j.smim.2016.05.004.

16. Kurtz, J.; Franz, K. Innate defence: evidence for memory in invertebrate immunity. Nature. 425, 37–38 (2003). doi:10.1038/425037a.

17. Sadd, B.M.; Schmid-Hempel, P. Insect immunity shows specificity in protection upon secondary pathogen exposure. Curr Biol. 16, 1206–1210 (2006). doi:10.1016/j.cub.2006.04.047.

18. Melillo, D.; Marino, R.; Italiani, P.; Boraschi, D. Innate Immune Memory in Invertebrate Metazoans: A Critical Appraisal. Front Immunol. 9, 1915 (2018). doi:10.3389/fimmu.2018.01915.

19. Portela, J.; Duval, D.; Rognon, A.; Galinier, R.; Boissier, J.; Coustau, C.; Mitta, G.; Théron, A.; Gourbal, B. Evidence for specific genotype-dependent immune priming in the lophotrochozoan *Biomphalaria glabrata* snail. J Innate Immun. 5, 261–276 (2013). doi:10.1159/000345909.

20. Cong, M.; Song, L.; Wang, L.; Zhao, J.; Qiu, L.; Li, L.; Zhang, H. The enhanced immune protection of Zhikong scallop *Chlamys farreri* on the secondary encounter with *Listonella anguillarum*. Comp Biochem Physiol B Biochem Mol Biol. 151, 191–196 (2008). doi:10.1016/j.cbpb.2008.06.014.

21. Wang, J.; Wang, L.; Yang, C.; Jiang, Q.; Zhang, H.; Yue, F.; Huang, M.; Sun, Z.; Song, L. The response of mRNA expression upon

secondary challenge with *Vibrio anguillarum* suggests the involvement of C-lectins in the immune priming of scallop *Chlamys farreri*. Dev Comp Immunol. 40, 142–147 (2013). doi:10.1016/j.dci.2013.02.003.

22. Pinaud, S.; Portela, J.; Duval, D.; Nowacki, F.C.; Olive, M.A.; Allienne, J.F.; Galinier, R.; Dheilly, N.M.; Kieffer-Jaquinod, S.; Mitta, G.; Théron, A.; Gourbal, B. A Shift from Cellular to Humoral Responses Contributes to Innate Immune Memory in the Vector Snail *Biomphalaria glabrata*. PLoS Pathog. 12, e1005361 (2016). doi:10.1371/journal.ppat.1005361.

23. Green, T.J.; Montagnani, C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus (OsHV-1 µvar). Fish Shellfish Immunol. 35, 382–8 (2013). doi:10.1016/j.fsi.2013.04.051.

24. Green, T.J.; Helbig, K.; Speck, P.; Raftos, D.A. Primed for success: oyster parents treated with poly(I:C) produce offspring with enhanced protection against Ostreid herpesvirus type I infection. Mol Immunol. 78, 113–20 (2016). doi:10.1016/j.molimm.2016.09.002.

25. Lafont, M.; Petton, B.; Vergnes, A.; Pauletto, M.; Segarra, A.; Gourbal, B.; Montagnani, C. Long-lasting antiviral innate immune priming in the Lophotrochozoan Pacific oyster, *Crassostrea gigas*. Sci Rep. 7, 13143 (2017). doi:10.1038/s41598-017-13564-0.

26. Romero, A.; Costa, M.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. 108, 149–163 (2014). doi:10.3354/dao02701.

27. Murgarella, M.; Puiu, D.; Novoa, B.; Figueras, A.; Posada, D.; Canchaya, C. A First insight into the genome of the filter-feeder mussel *Mytilus galloprovincialis*. PLoS ONE. 11, e0151561 (2016). doi:10.1371/journal.pone.0151561.

28. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31, 3210–2 (2015). doi:10.1093/bioinformatics/btv351.

29. McCarthy, D.J.; Chen, Y.; Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 40, 4288–97 (2012). doi:10.1093/nar/gks042.

30. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21, 3674–6 (2005). doi:10.1093/bioinformatics/bti610.

31. Costa, M.M.; Novoa, B.; Figueras, A. Influence of b-glucans on the immune responses of carpet shell clam (*Ruditapes decussatus*) and Mediterranean mussel (*Mytilus galloprovincialis*). Fish Shellfish Immunol. 24, 498–505 (2008). doi:10.1016/j.fsi.2007.10.003.

32. Netea, M.G.; Quintin, J.; Van der Meer, J.W.M. Trained immunity: a memory for innate host defense. Cell Host Microbe. 9, 355–61 (2011). doi:10.1016/j.chom.2011.04.006.

33. Quintin, J.; Cheng, S.-C.; Van der Meer, J.W.M.; Netea, M.G. Innate immune memory: towards a better understanding of host defense mechanisms. Curr Opin Immunol. 29, 1–7 (2014). doi:10.1016/j.coi.2014.02.006.

34. Gourbal, B.; Pinaud, S.; Beckers, G.J.M.; Van der Meer, J.W.M.; Conrath, U.; Netea, M.G. Innate immune memory: an evolutionary perspective. Immunol Rev. 283, 21–40 (2018). doi:10.1111/imr.12647.

35. Green, T.J.; Benkendorff, K.; Robinson, N.; Raftos, D.; Speck, P. Anti-viral gene induction is absent upon secondary challenge with double-stranded RNA in the Pacific oyster, *Crassostrea gigas*. Fish Shellfish Immunol. 39, 492–7 (2014). doi:10.1016/j.fsi.2014.06.010.

36. Ivanenkov, Y.A.; Balakin, K.V.; Lavrovsky, Y. Small molecule inhibitors of NF-kB and JAK/STAT signal transduction pathways as promising anti-inflammatory therapeutics. Mini Rev Med Chem. 11, 55–78 (2011). doi:10.2174/138955711793564079.

37. Nathan, C. Points of control in inflammation. Nature. 420, 846–52 (2002). doi:10.1038/nature01320.

38. Tobin, A.E.; Calabrese, R.L. Myomodulin increases Ih and inhibits the NA/K pump to modulate bursting in leech heart interneurons. J Neurophysiol. 94, 3938–50 (2005). doi:10.1152/jn.00340.2005.

39. Eisenhut, M. Changes in ion transport in inflammatory disease. J Inflamm. 3, 5 (2006). doi:10.1186/1476-9255-3-5.

40. Kravtsova-Ivantsiv, Y.; Kwon, Y.T.; Ciechanover, A. Role of the ubiquitin ligase KPC1 in NF-kB activation and tumor suppression. J Anal Sci Technol. 7, 8 (2016). doi:10.1186/s40543-016-0087-4.

41. McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scanu, A.M.; Lawn, R.M. cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. Nature. 330, 132–7 (1987). doi:10.1038/330132a0.

42. Ichinose, A. Characterization of the apolipoprotein(a) gene. Biochem Biophys Res Commun. 209, 365–71 (1995). doi:10.1006/bbrc.1995.1512.

43. Sulniute, R.; Shen, Y.; Guo, Y.Z.; Fallah, M.; Ahlskog, N.; Ny, L.; Rakhimova, O.; Broden, J.; Boija, H.; Moghaddam, A.; Li, J.; Wilczynska, M.; Ny, T. Plasminogen is a critical regulator of cutaneous wound healing. Thromb Haemost. 115, 1001–9 (2016). doi:10.1160/TH15-08-0653.

44. Zha, J.; Zhou, Q.; Xu, L.G.; Chen, D.; Li, L.; Zhai, Z.; Shu, H.B. RIP5 is a RIP-homologous inducer of cell death. Biochem Biophys Res Commun. 319, 298–303 (2004). doi:10.1016/j.bbrc.2004.04.194.

45. Manduzio, H.; Rocher, B.; Durand, F.; Galap, C.; Leboulenger, F. The point about oxidative stress in molluscs. Invert Surviv J. 2, 91–104 (2005).

46. Bartosz, G. Reactive oxygen species: destroyers or messengers? Biochem Pharmacol. 77, 1303–15 (2009). doi:10.1016/j.bcp.2008.11.009.

47. Donaghy, L.; Lambert, C.; Choi, K.S.; Soudant, P. Hemocytes of the carpet shell clam (*Ruditapes decussatus*) and the Manila clam (*Ruditapes philippinarum*): current knowledge and future prospects. Aquaculture. 297, 10–24 (2009). doi:10.1016/j.aquaculture.2009.09.003.

48. Brand, M.D.; Esteves, T.C. Physiological functions of the mitochondrial uncoupling proteins UCP2 and UCP3. Cell Metab. 2, 85–93 (2005). doi:10.1016/j.cmet.2005.06.002.

49. Li, R.; Chen, W.; Yanes, R.; Lee, S.; Berliner, J.A. OKL38 is an oxidative stress response gene stimulated by oxidized phospholipids. J Lipid Res. 48, 709–15 (2007). doi:10.1194/jlr.M600501-JLR200.

50. Ying, W. NAD+/NADH and NADP+/NADPH in cellular functions and cell death: regulation and biological consequences. Antioxid Redox Signal. 10, 179–206 (2008). doi:10.1089/ars.2007.1672.

51. Bednarski, J.J.; Sleckman, B.P. At the intersection of DNA damage and immune responses. Nat Rev Immunol. 19, 231–42 (2019). doi:10.1038/s41577-019-0135-6.

52. Sugasawa, K. Xeroderma pigmentosum genes: functions inside and outside DNA repair. Carcinogenesis. 29, 455–65 (2008). doi:10.1093/carcin/bgm282.

53. Simon, H.U.; Haj-Yehia, A.; Levi-Schaffer, F. Role of reactive oxygen species (ROS) in apoptosis induction. Apoptosis. 5, 415–8 (2000). doi:10.1023/A:1009616228304.

54. Fleury, C.; Mignotte, B.; Vayssière, J.L. Mitochondrial reactive oxygen species in cell death signaling. Biochimie. 84, 131–41 (2002). doi:10.1016/S0300-9084(02)01369-X.

55. Jia, J.; Furlan, A.; Gonzalez-Hilarion, S.; Leroy, C.; Gruenert, D.C.; Tulasne, D.; Lejeune, F. Caspases shutdown nonsense-mediated

mRNA decay during apoptosis. Cell Death Differ. 22, 1754–63 (2015). doi:10.1038/cdd. 2015.18.

56. Kapodistria, K.; Tsilibary, E.P.; Politis, P.; Moustardas, P.; Charonis, A.; Kitsiou, P. Nephrin, a transmembrane protein, is involved in pancreatic beta-cell survival signaling. Mol Cell Endocrinol. 400, 112–28 (2015). doi:10.1016/j.mce.2014.11.003.

57. Lin, P.; Mobasher, M.E.; Alawi, F. Acute dyskerin depletion triggers cellular senescence and renders osteosarcoma cells resistant to genotoxic stress-induced apoptosis. Biochem Biophys Res Commun. 446, 1268–75 (2014). doi:10.1016/j.bbrc.2014. 03.114.

58. Zhou, T.; Zimmerman, W.; Liu, X.; Erikson, R.L. A mammalian NudC-like protein essential for dynein stability and cell viability. Proc Natl Acad Sci USA. 103, 9039–44 (2006). doi:10.1073/pnas.0602916103.

59. Cong, M.; Song, L.; Wang, L.; Zhao, J.; Qiu, L.; Li, L.; Zhang, H. The enhanced immune protection of Zhikong scallop *Chlamys farreri* on the secondary encounter with *Listonella anguillarum*. Comp Biochem Physiol B Biochem Mol Biol. 151, 191–6 (2008). doi:10.1016/j.cbpb.2008.06.014.

60. Zhang, T.; Qiu, L.; Sun, Z.; Wang, L.; Zhou, Z.; Liu, R.; Yue, F.; Sun, R.; Song, L. The specifically enhanced cellular immune responses in Pacific oyster (*Crassostrea gigas*) against secondary challenge with *Vibrio splendidus*. Dev Comp Immunol. 45, 141–50 (2014). doi:10.1016/j.dci.2014.02.015.

61. Pinaud, S.; Portela, J.; Duval, D.; Nowacki, F.C.; Olive, M.A.; Allienne, J.F.; Galinier, R.; Dheilly, N.M.; Kieffer-Jaquinod, S.; Mitta, G.; Théron, A.; Gourbal, B. A shift from cellular to humoral responses contributes to innate immune memory in the vector snail *Biomphalaria glabrata*. PLoS Pathog. 12, e1005361 (2016). doi:10.1371/journal.ppat.1005361.

62. Dubief, B.; Nunes, F.L.D.; Basuyaux, O.; Paillard, C. Immune priming and portal of entry effectors improve response to *Vibrio* infection in a resistant population of the European abalone. Fish

Shellfish Immunol. 60, 255–64 (2017). doi:10.1016/j.fsi.2016.11.017.

63. Beets, I.; Janssen, T.; Meelkop, E.; Temmerman, L.; Suetens, N.; Rademakers, S.; Jansen, G.; Schoofs, L. Vasopressin/oxytocin-related signaling regulates gustatory associative learning in *C. elegans*. Science. 338, 543–5 (2012). doi:10.1126/science.1226860.

64. Hommes, F.A. Loss of neurotransmitter receptors by hyperphenylalaninemia in the HPH-5 mouse brain. Acta Paediatr Suppl. 407, 120–1 (1994).

65. Gassió, R.; Vilaseca, M.A.; Lambruschini, N.; Boix, C.; Fusté, M.E.; Campistol, J. Cognitive functions in patients with phenylketonuria in long-term treatment with tetrahydrobiopterin. Mol Genet Metab. 99 Suppl. 1, S75–78 (2010). doi:10.1016/j.ymgme.2009.10.187.

66. Mitchell, J.J.; Trakadis, Y.J.; Scriver, C.R. Phenylalanine hydroxylase deficiency. Genet Med. 13, 697–707 (2011). doi:10.1097/GIM.0b013e3182141b48.

CHAPTER 4

TRANSCRIPTOMIC ANALYSIS REVEALS THE WOUND HEALING ACTIVITY OF MUSSEL MYTICIN C

Rey-Campos, M.[1]; Moreira, R.[1]; Romero, A.[1]; Medina-Gali, R.M.[2]; Novoa, B.[1]; Gasset, M.[3]; Figueras, A.[1]. Transcriptomic analysis reveals the wound healing activity of mussel myticin C. Biomolecules. 10, 133 (2020). doi:10.3390/biom10010133. (Open Access).

[1]Institute of Marine Research (IIM), National Research Council (CSIC), Eduardo Cabello, 6, 36208, Vigo, Spain. [2]Instituto de Investigación, Desarrollo e Innovación en Biotecnología Sanitaria de Elche (IDiBE) and Instituto de Biología Molecular y Celular (IBMC), Miguel Hernández University (UMH), 03202 Elche Alicante, Spain. [3]Instituto Química-Física "Rocasolano", CSIC. Serrano 119, 28006 Madrid, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 4: Transcriptomic analysis reveals the wound healing activity of mussel myticin C

4.1 INTRODUCTION

Although Mediterranean mussels (*Mytilus galloprovincialis*) are mainly known for their economic relevance in aquaculture [1], it is also interesting because of their singular resistance to diseases, that associated to the lack of an adaptive immune system, make mussels an interesting model to study immune molecules, especially antimicrobial peptides (AMPs).

These animals are characterized by an open circulatory system, in which circulating hemocytes would be primarily responsible for the immune response. Hemocytes are capable of encapsulating and phagocyting foreign particles or pathogens, as well as synthesizing and releasing cytotoxic factors [2], including AMPs that represent key components of the mussel immune system. These small cationic peptides are cysteine-rich molecules, and are traditionally involved in the response to bacteria, some fungi and viruses [3,4].

So far, nine AMPs have been identified in mussels: defensins [5], myticins [6], mytilins [7], mytimycins [8], big defensins [9], mytimacins [9], myticusins [10], mytichitins [11], and myticalins [12]. Some of these AMPs exhibit an extraordinary diversity in their structure and function, which could be related to a more specific and improved defense against pathogens. In the case of myticins, three isoforms have been characterized (A, B, and C) [6,13]. Myticin C has been the most studied myticin so far. It shows a high variability at nucleotide level,

even displaying an exclusive repertoire in each individual mussel [14]. How this variability is generated, and what role it plays in the mussel immune response, is a field yet to be explored and which will be the subject of the following chapter of this thesis.

Myticin C is constitutively expressed in mussel hemocytes and stored in vesicles in the cytoplasm [4]. From a functional point of view, myticin C shows antibacterial activity [2] and antiviral function against fish rhabdovirus [15], ostreid herpesvirus (OsHV-1), and even human herpes simplex (HSV-1 and HSV-2) [4]. Recently, apart from the antimicrobial activity, a new function related to damage-associated molecular pattern (DAMP) response and response to tissue injury has been suggested (Chapter 2 of this thesis). Moreover, chemotactic properties have been attributed to myticin C [15]. This capacity to promote hemocyte migration makes myticin C a chemokine-like molecule, and therefore, its function triggers cellular mechanisms like adhesion, spreading, migration, and phagocytosis.

The main objective of this work was to characterize the hemocytes' transcriptomic response after a myticin C treatment, in order to understand the molecular changes implied in the response to this peptide. The ability to attract hemocytes with high myticin C concentration towards damaged tissue has allowed us to investigate the potential biotechnological application of myticin C in the wound healing of vertebrates, including humans.

4.2 MATERIALS AND METHODS

4.2.1 Animals

Adult *M. galloprovincialis*, 6–8 cm in shell length, were obtained from a commercial shellfish farm (Vigo, Galicia, Spain) and maintained in open-circuit, filtered sea water tanks at 15 ºC with aeration. The animals were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, the animals were acclimatized to aquarium conditions for one week.

The Mediterranean mussel, *M. galloprovincialis*, is not considered an endangered or protected species in any international species catalogue, including the Convention on International Trade in Endangered Species (CITES) list (www.cites.org). *M. galloprovincialis* is not included in the European Union (EU) regulation to work with research animals by the European Directive 2010/63/EU. Therefore, no specific authorization is required to work with these samples.

Wild-type zebrafish larvae were obtained from the facilities at the Instituto de Investigaciones Marinas (Vigo, Spain), where zebrafish are maintained following established protocols. Zebrafish were euthanized using a tricaine methanesulfonate (MS-222) overdose (500 mg/L). Fish care and regeneration experiments were conducted according to the guidelines of the CSIC (Spanish National Research Council—Consejo Superior de Investigaciones Científicas) National Committee on Bioethics, under approval number ES360570202001/16/FUN01/PAT.05/tipoE/BNG.

4.2.2 Transcriptomic experimental approach

Mussels were notched in the shell, and 1 mL of hemolymph was withdrawn from the adductor muscle of each mussel with a 0.5 mm diameter (25G) disposable needle. Hemolymph was pooled (three pools of 25 mussels) and placed in a six-well polystyrene plate (BD, Falcon; 5 mL per pool) for 30 min at 15 ºC to let it settle. Hemocytes were then treated with myticin C. The synthetic myticin C mature peptide [16] was manufactured by GenScript (Leiden, Netherlands) with a purity >95%, determined by high-performance liquid chromatography and mass spectrometry. The final concentration of myticin C was 10 µM (control hemocytes remained unstimulated). This experimental condition was maintained for 8 h at 15 ºC. Sampling was performed by scraping the hemocytes from the bottom of the well. Hemocytes were centrifuged at 4 ºC at 3000 x g for 10 min, and the pellet was then resuspended in 300 µl of homogenation buffer (Promega, Madison, WI; USA) and immediately homogenized with a syringe and needle. RNA isolation was carried out using the Maxwell 16 LEV robot, following the instructions for the simplyRNA kit (Promega). Next, the

concentration and purity of the RNA was measured using a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA), and RNA integrity was tested on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) before producing the sequencing libraries.

An Illumina TruSeq Stranded mRNA LT Sample Preparation Kit (San Diego, CA, USA) was used, according to the manufacturer's instructions. Briefly, eukaryotic mRNA was extracted from total RNA using oligo (dT) magnetic beads, and was cleaved into short fragments using a 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, Hilden; Germany). The amount of cDNA in each library was quantified through spectrofluorometric analysis, using the Qubit system. Next-generation sequencing was performed using Illumina HiSeq 4000 technology in Macrogen (Seoul, Korea). The raw reads (101 nucleotides) were deposited in the NCBI (National Center for Biotechnology Information) database with the following accession numbers: SAMN09104581, SAMN09104582, and SAMN09104583 for the control samples; and SAMN09104593, SAMN09104594, and SAMN09104595 for myticin C-treated samples.

4.2.3 Bioinformatics: assembly, RNA-Seq, and annotation

CLC GenomicsWorkbench, v.11.0.1 [17], was used to trim, assemble, and perform the RNA-seq and statistical analysis. Raw reads were trimmed to remove adaptor sequences, low-quality sequences (PHRED = 13), and sequences less than 70 bp. Then a reference global transcriptome of the six libraries was assembled, with a minimum contig length of 200 bp. Next RNA-seq analysis was performed, with default settings, to obtain TPM (Transcripts Per Million) expression values. To identify differentially expressed genes (DEGs), a Robinson and Smyth's Exact Test was carried out [18]. Transcripts with absolute fold change (FC) values >2 and a false discovery rate (FDR)-corrected

p-value <0.05 were retained for further analyses. Blast2GO software [19] was used to obtain UniProt/Swiss-Prot annotations and gene ontology (GO) term assignments for the contig list. A BLASTn approach was also performed, with an in-house built database made with all the mollusc sequences present in the NBCI nucleotide database. The e-value threshold was set at 1e-3. Then, an enrichment analysis of DEGs (test set) was conducted, including the global hemocyte transcriptome as the reference set. A Fisher's exact test [20] was run with a false discovery rate (FDR) cut-off of 0.05. The option to show only the most specific terms was used. Over-represented biological processes (BPs), molecular functions (MFs), and cellular components (CCs) were further analyzed. Finally, Blast2GO was also used to analyze KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways in which DEGs were involved.

4.2.4 Hemocyte time-lapse microscopy and morphological analysis

Hemolymph extracted from the adductor muscle was diluted 1:20 in filtered sea water (FSW) and distributed in a 24-well polystyrene plate (BD, Falcon), 500 µl per well, and incubated 30 min at 15 ºC to let the hemocytes settle and adhere. Two parallel plates were prepared with the same hemolymph. The mean concentration of hemocytes was 10^5 cells/mL. Hemocytes were stimulated with a solution containing myticin C at a final concentration of 20 µM or with FSW (control cells). Control hemocytes were imaged on a Nikon TMS inverted microscope equipped with phase contrast objectives and a Nikon DMX 1200 camera (Tokyo, Japan). Greyscale images (3840 x 3072 pixels) were acquired using the Nikon ACT-1 v2.7 acquisition software and calibrated (1 px = 0.1366 µm). Stimulated hemocytes were observed in the TS100 Eclipse inverted microscope (Nikon) equipped with a DS-Fi1 camera (Nikon). Eight-bit images (2569 x 1920 pixels) were acquired using the Nis-Elements V2.32 software and calibrated (1 px = 0.33 µm) (Nikon). Time-lapse recordings were performed in both plates in parallel, capturing images every 30 s for 3 h of stimulation. Individual images of the time-lapse sequence were processed to

compute morphological and movement parameters, using manual tracking and the Chemotaxis and Migration Tool 2.0 plugins for the ImageJ analysis software [21]. Three independent stimulations were conducted using hemolymph from single animals. In each experiment, at least 40 individual cells were analyzed, and the maximum cell length, area, mean velocity, and accumulated distance were measured. One-way ANOVA with post-hoc Tukey test was conducted using GraphPad Prism software (san Diego; CA; USA), and results were considered significant, with a threshold p-value < 0.05.

The size and morphology of hemocytes were also analyzed in fixed samples. Briefly, after stimulation, cells were fixed in 4% paraformaldehyde (PFA) and permeabilized in 1% triton X-100/PBS for 3 min. Unspecific sites were blocked with 1% bovine serum albumin (BSA) overnight at room temperature. Next, cells were stained with 0.165 µM rhodamine-phalloidin (Molecular Probes, Invitrogen, Carlsbad, CA, USA) and 0.1 µg/mL 4',6-diamidino-2-phenylindole (DAPI; Molecular Probes, Invitrogen, Carlsbad, CA, USA). Samples were mounted on slides using ProLong Gold (Molecular Probes, Invitrogen) reagent and visualized on an TCS SPE fluorescent microscope (Leica, Wetzlar, Germany).

4.2.5 Histological and immunofluorescence assays

Nine mussels were notched in the shell. Three mussels were injured in the adductor muscle using a 21G disposable needle, another three mussels were stimulated with myticin C (10 µM), and the last three mussels were treated as controls (injected with FSW). Also, three naïve mussels were included in the experiment. Animals were maintained in 10 L tanks at 15 ºC with aeration for 4 h before sampling. Posterior adductor muscles were extracted, immediately fixed for 24 h in 9/1 Davidson solution/acetic acid, and then stored in Davidson solution until the sample was embedded in paraffin. Histological sections were stained with hematoxylin and eosin (Merck, Kenilworth, NJ, USA) and examined under light microscopy (Nikon Eclipse 80i).

Histological sections of 4 µm were also used for an immunofluorescence assay. Fixed muscles were incubated overnight (4 ºC) with a rabbit polyclonal anti-myticin C antibody (1:50) [20] and a mouse monoclonal anti-actin antibody (1:100) (Clon C4, Millipore, Burlington, MA, USA). Alexa Fluor 546-conjugated anti-rabbit and Alexa Fluor 488-conjugated anti-mouse (1:500 and 1:1000, respectively; Life Technologies, Carlsbad, CA, USA) were used as secondary antibodies. The slides were stained with DAPI (Molecular Probes, Invitrogen, Carlsbad, CA, USA) and mounted using ProLong antifade reagents (Life Technologies, Carlsbad, CA, USA). The images were captured using a TSC SPE confocal microscope (Leica, Wetzlar, Germany) and processed using LAS-AF (Leica) and ImageJ software. The same experimental approach was replicated to obtain hemolymph samples and count hemocytes in a Neubauer chamber. These counts were performed 4 h and 24 h after tissue injury.

4.2.6 Western blot of myticin C

Four pools of three mussels were injured in the adductor muscle using a 21G disposable needle (the same number of mussels were treated as controls). Animals were maintained in 10 L tanks at 15 ºC with aeration for 4 h before sampling, and hemolymph (1 mL per mussel) was withdrawn from the adductor muscle. After that, hemocytes were separated by centrifugation (3,000g, 10 min). These cells were lysed with 100 µL of ice-cold lysis buffer (50 mM Tris-HCl pH 7.8, 0.25 M sucrose, 1% SDS (Sodium Dodecyl Sulfate), 5 mM EDTA (Ethylenediaminetetraacetic Acid), 0.1% Nonidet-P40) with a 1% protease inhibitor cocktail and a phosphatase inhibitor cocktail (Sigma-Aldrich, St. Louis, MO, USA, catalog numbers P8849 and 78420, respectively), and then centrifuged for 15 min at 14,000 rpm to remove insoluble debris.

All samples (normalized to 50 µg of protein before loading) were mixed 3:1 (v:v) with the sample buffer (277.8 mM Tris-HCl pH 6.8, 4.4% LDS, 44.4% glycerol, 10% 2-mercaptoethanol, 0.02% Bromophenol blue) and boiled at 95 ºC for 5 min. After that, samples were resolved by SDS-polyacrylamide gel electrophoresis on 4–20%

Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad, Hercules, CA, USAes). Molecular mass marker Precision Plus Protein Dual Color Standards (Bio-Rad) were run on adjacent lanes. The gels were electro-blotted in nitrocellulose membranes (45 min of blotting time), and the blots were probed with primary antibodies: rabbit polyclonal anti-myticin C antibody (1:500, overnight at 4 ºC) [15] and a mouse monoclonal anti-actin antibody (1:5000, 1 h at room temperature) (Clon C4, Millipore, Burlington, MA, United States). Peroxidase-conjugated goat anti-rabbit IgG (1:6000) (A6154; Sigma-Aldrich, St. Louis, MO, United States) and anti-mouse IgG (1:8000) (A4416; Sigma-Aldrich, St. Louis, MO, USA) were used as secondary antibodies. Membranes were visualized using an enhanced chemiluminescenceWestern blotting analysis system (Immobilon Forte Western HRP Substrate), following the manufacturer's specifications. Western blot films were digitized (Chemidoc XRS+, Bio-Rad), and band optical densities (arbitrary units) were quantified using a computerized imaging system (Image Lab, Bio-Rad).

4.2.7 Wound healing and tail fin regeneration assays

HaCaT cells (immortalized human keratinocyte line) (CLS cell lines service, Germany, Cat nº 300493) were cultured in Dulbecco's Modified Eagle Medium (Gibco, Gaithersburg, MD, USA), supplemented with 10% fetal bovine serum (Gibco) and 50 µg/mL of gentamicin (Gibco) at 37 ºC and 5% CO2. HaCaT monolayers on a 24-well plate were scratched and imaged on a TS100 Eclipse inverted microscope (Nikon) equipped with a DS-1QM camera (Nikon). After that, HaCaT cells were treated with the synthetic myticin C at a final concentration of 5 µM. Images were taken at 24, 48, and 72 h, and the area of opening was measured using ImageJ software. For each photograph, 10 measures of the open wound were recorded, averaged, and compared with the control samples. GraphPad Prism software was used for statistical analysis, performed by unpaired t test with Welch's correction.

Groups of 20 wild-type (WT) zebrafish larvae (three days post-fertilization) were anesthetized with MS-222, submitted to a tail fin cut,

and exposed to the peptide at a final concentration of 5 µM. All zebrafish larvae were imaged using a Nikon AZ100 lens equipped with a DS-Fi1 camera (Nikon). Measures of the tail fins at day 0, 2, 4, and 7 were performed, using the posterior notochord as a reference to the end of the tail. Again, GraphPad Prism software was used for statistical analysis, performed by unpaired t test with Welch's correction.

4.3 RESULTS

4.3.1 Assembly and annotation of hemocyte transcriptome

The sequencing of hemocyte samples yielded an average of 72.6 million raw reads. Raw reads were imported into the CLC Genomics Workbench for trimming, assembly, and gene expression analysis. Quality control trimmed out 1.57% of raw reads. The remaining 98.43%, corresponding to the six individual libraries, were assembled into a global mussel transcriptome containing 154,093 contigs, with an average length of 509 bp. Then, two different BLAST approaches were followed to identify these contigs: CLC was used to identify 31.72% of the contigs, using an in-house designed database with all the nucleotidic sequences available in NCBI for molluscs, and Blast2GO was used to identify 19.86% of the contigs through BLASTx against the UniProt/Swiss-Prot database. Gene ontology (GO) terms were assigned to 19.74% of the contigs, and 9,118 contigs were present in seven KEGG pathways. This information is shown in Table 1.

Table 1. Summary of the transcriptome bioinformatics results.

Reads Origin	Raw reads	Trimmed reads
Control 1	90,967,138	99.53%
Control 2	84,688,626	99.49%
Control 3	53,800,576	94.49%
MytC 1	73,793,348	99.42%
MytC 2	80,981,306	98.79%
MytC 3	51,346,136	98.88%
Assembly		
Contigs	154,093	
Range contig length	200–16,293	
Average contig length	509	
N50	568	
Blast		
Contigs identified by UniProt/Swiss-Prot	30,596	
Contigs identified by molluscs database	48,876	
GO analysis		
Annotated contigs	30,416	
KEGG analysis		
Pathway assigned contigs	9,118	

4.3.2 Hemocyte transcriptomic response after a myticin C treatment

The RNA-seq analysis of stimulated hemocytes revealed 45 differentially expressed genes (DEGs) -39 up-regulated and 6 down-regulated- after a myticin C treatment (available in https://www.mdpi.com/2218-273X/10/1/133/s1). Enriched GO terms after Fisher's exact test showed that the most specific terms (Figure 1a) were related to the cytoskeleton and contraction, as well as maintenance of integrity of tissues and cell structures ("actin-dependent ATP-ase activity", "microtubule binding", or "actin filament binding"). Additionally, KEGG pathways were also represented (Figure 1b), with the pathways related to nucleic acid metabolism (purine and pyrimidine metabolism pathways) and vertebrate immune systems ("T cell receptor signaling pathway" and "Th1 and Th2 cell differentiation") being the most represented.

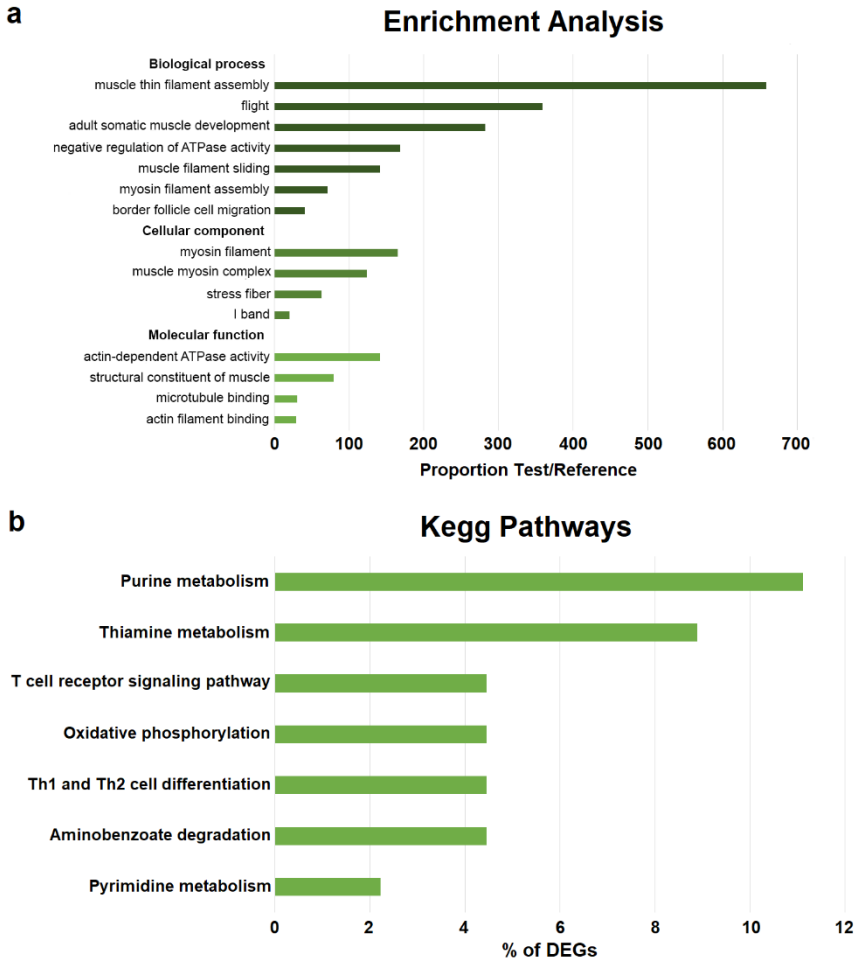

Figure 1. (a) Enrichment analysis of differentially expressed genes (DEGs). Bars represent the significant changes of proportion between the percentage of sequences in the DEG list and the transcriptome. **(b)** KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways related to the regulated transcripts after a myticin C treatment.

All the annotated DEGs are reported in Table 2, with their respective fold changes. Myosin is one of the most represented and up-regulated contigs in the transcriptome (over two folds, compared to controls). Myosin provides the motor function for diverse movements, such as cytokinesis, phagocytosis, and muscle contraction. Cellular

mobility is characterized by podosome formation. These structures are actin-rich, and their formation is influenced by genes belonging to the family of calponins (transgelin-like proteins or calponin-like proteins), which are also up-regulated. In the same way, RS-rich proteins and DCST2 are over expressed in hemocytes after treatment. These proteins are involved in processes related to the cellular cycle, cell structure, and mobility. Finally, HSPG2 (Heparan Sulfate Proteoglycan 2) is a glycoprotein involved in adhesion, migration, and differentiation through the mediation of cell adhesion molecules. This gene product is named "perlecan" and is a multifunctional proteoglycan that preserves the integrity of extracellular matrices.

Table 2. Top annotated DEGs after a myticin C treatment. FC, fold change.

Contig	FC	Description
Mg_121683	59.81	Cytochrome c oxidase subunit 1
Mg_151173	14.74	Cytoglobin-1
Mg_40784	7.55	LKD-rich protein-1
Mg_135551	5.87	Vitelline envelope zona pellucida domain 9
Mg_53155	4.97	Heparan Sulfate Proteoglycan 2 (HSPG2)
Mg_137132	4.85	Vitelline envelope zona pellucida domain 9
Mg_73904	4.56	DnaJ homolog subfamily B member 5 (DNAJB5/HSP40)
Mg_135670	4.31	Serine protease inhibitor dipetalogastin
Mg_99492	4.22	Tripartite motif-containing protein 56 (TRIM56)
Mg_136869	4.21	Serine protease inhibitor dipetalogastin
Mg_74563	4.02	Sarcoplasmic calcium-binding protein
Mg_34823	3.62	RS-rich protein-1
Mg_34824	3.50	RS-rich protein-2
Mg_35853	3.26	Transgelin-like protein-6
Mg_22493	3.23	DC-STAMP domain-containing protein 2 (DCST2)
Mg_12073	3.07	Myosin heavy chain
Mg_65743	3.04	RS-rich protein-1
Mg_30352	2.95	Calponin-like protein
Mg_90861	2.93	LKD-rich protein-1
Mg_60714	2.88	Myosin heavy chain
Mg_40783	2.83	LKD-rich protein-1
Mg_25094	2.76	Protein SOGA3
Mg_19629	2.70	Calponin-like protein
Mg_268	2.69	Myosin regulatory light chain
Mg_18832	2.57	Small heat shock protein 22
Mg_17930	2.57	Nicotinamidase
Mg_31755	2.49	RS-rich protein-2
Mg_49773	2.46	Myosin
Mg_34500	2.29	DBH-like monooxygenase protein 1 homolog

Genes related to O_2 homeostasis (cytochrome c oxidase subunit 1 and cytoglobin) or the regulation of cell physiology under normal and stress conditions (HSP22 and HSP40) are also up-regulated in the myticin C-treated hemocyte transcriptome.

4.3.3 Changes in hemocytes induced by myticin C (morphology, mobility, and number)

To observe the effect that myticin C exerts on hemocytes, these cells were exposed to the peptide and then stained with phalloidin (Figure 2a). The formation of actin structures that accumulate in cytoplasm and their numerous extensions to the outside of the cells were observed. These structures could be reminiscent of podosomes, which are involved in cell mobility. Moreover, changes in actin levels were also observed *in vivo* after treating mussels with myticin C. An immunofluorescence assay on fixed adductor muscles was performed, and an increase in the actin levels after myticin C treatment could be observed (Figure 2b).

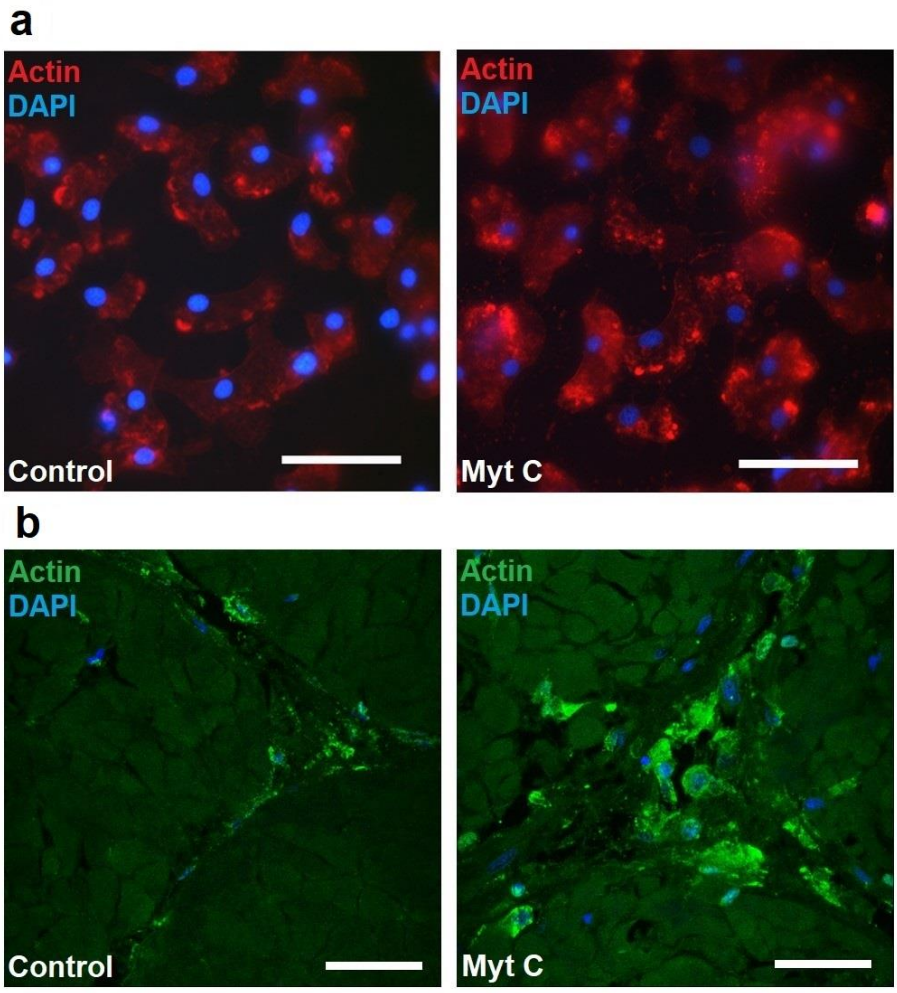

Figure 2. (a) Differential interference contrast (DIC) images of live hemocytes 3 h after stimulation with myticin C. Actin is stained in red and nuclei in blue. Scale bar = 25 µm. **(b)** Immunofluorescence of mussel adductor muscle. Actin is stained in green and nuclei in blue. Scale bar = 25 µm.

Morphological changes in hemocytes after myticin C treatment were evaluated by measuring the maximum cell length and area on differential interference contrast (DIC) microscopic images. The treatment with myticin C induced morphological changes in cells. At the beginning of the experiment, control and stimulated hemocytes

showed extended cytoplasms attached to the surface, and the mean size of hemocytes was 33.2±1.4 µm. Treatment with myticin C induced a significant increment of cell size and area. After 1 h of stimulation, hemocytes were 18.5% longer than controls (40.5±1.2 µm in stimulated vs. 34.2±1.4 µm in controls). The highest significant difference in size was at 2 h after stimulation, when treated cells were 36% longer than controls (41.7±1.3 µm in stimulated vs. 30.6±1.4 µm in controls) (Figure 3a).

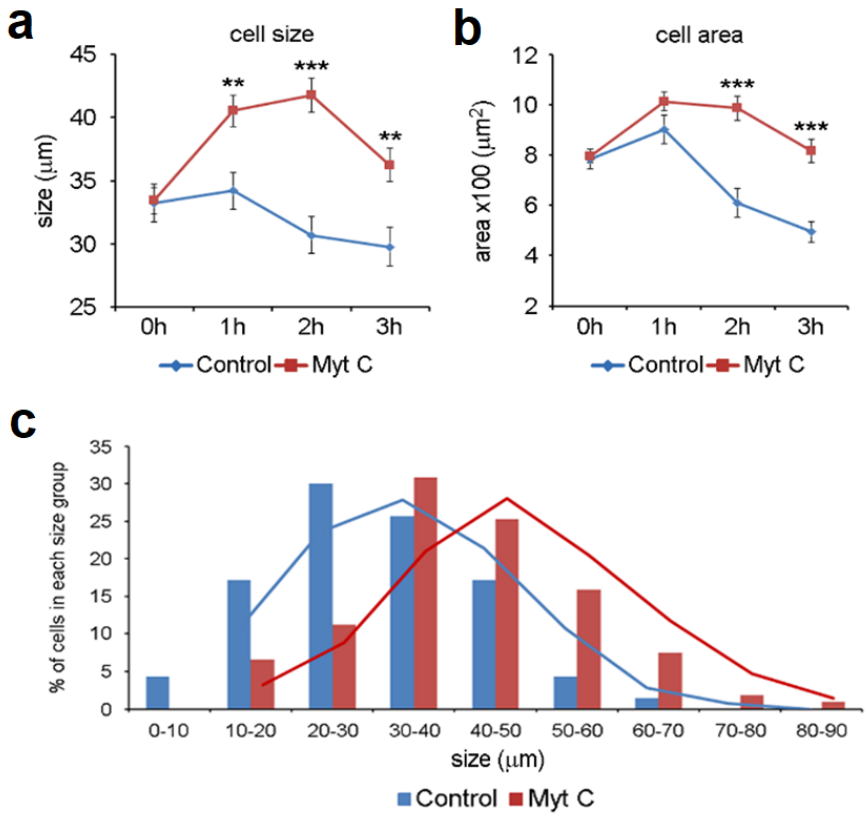

Figure 3. (a) Cell size analysis based on maximum length. **(b)** Cell size analysis based on area measurements. Asterisks indicates significant differences at $p < 0.01$ (**) and $p < 0.001$ (***). **(c)** Distribution of cells in different size groups after 3 h of stimulation.

The increment of cell size reflected a significant increment in cell area. Treated cells were 61.7% and 64.8% greater in area than controls at 2 and 3 h after stimulation, respectively (Figure 3b). Myticin C treatment changed the distribution of hemocytes in different cell size groups. The percentage of cells included in the bigger size groups (40–50, 50–60, and 60–70 μm) increased after 2 and 3 h post-treatment (Figure 3c).

The time-lapse recording generated a total of 361 sequential images during 3 h of stimulation. For the analysis, only individual cells that could be tracked throughout the entire experiment were selected. In this representative experiment, a total of 40 control cells and 72 stimulated cells were used (Figure 4a). The movement of cells is represented by a line. Cell velocity and accumulated distance were measured. Although myticin C did not modify the mean velocity of hemocytes (Figure 4b), the stimulated cells were able to travel longer distances than controls, yielding an increased accumulated distance (Figure 4c).

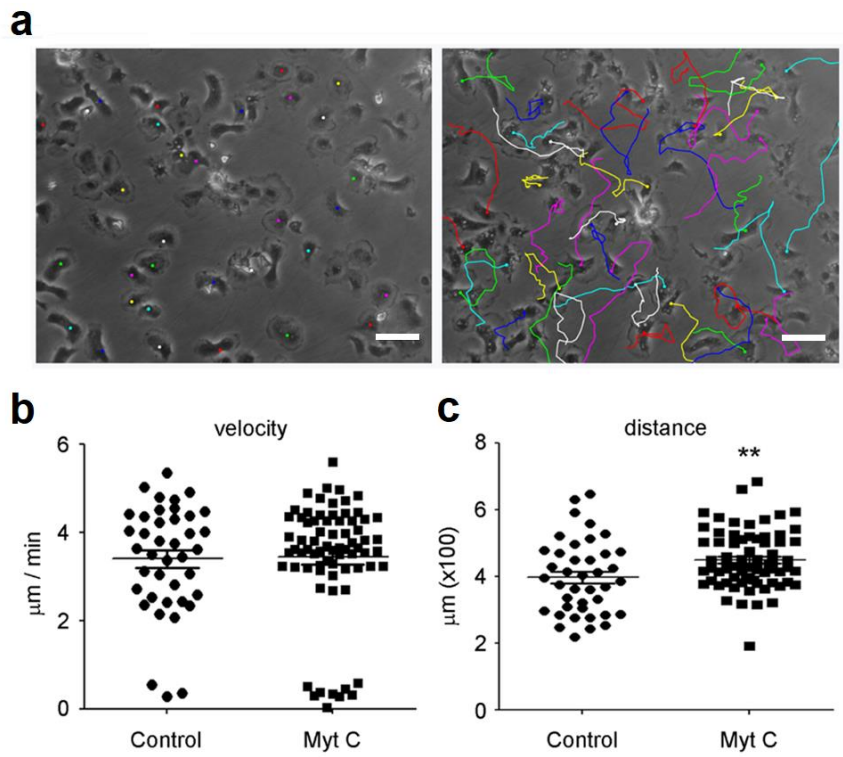

Figure 4. (a) Tracking of individual cell movements. The displacement of single cells is marked with a different color. Scale bar 50 µm. **(b)** Mean velocity of cell displacement (µm/min) after 3 h of stimulation. **(c)** Accumulated distance traveled by the cells during 3 h. Asterisks denote significant differences at $p < 0.01$ (**).

4.3.4 Effect of myticin C on injured tissues

Tissue injury triggers a cascade of expression of different genes and factors, with the aim of repairing the damaged tissue. This process is usually initiated by the migration of cells in charge of repairing the damage. An injury in the adductor muscle of mussels induced a significant increase in the number of hemocytes (in both times assessed, 4 h and 24 h) in the damaged area (Figure 5a). Moreover, figure 5b contains an image of a histological preparation, where the wound and the migration of hemocytes to the area can be observed.

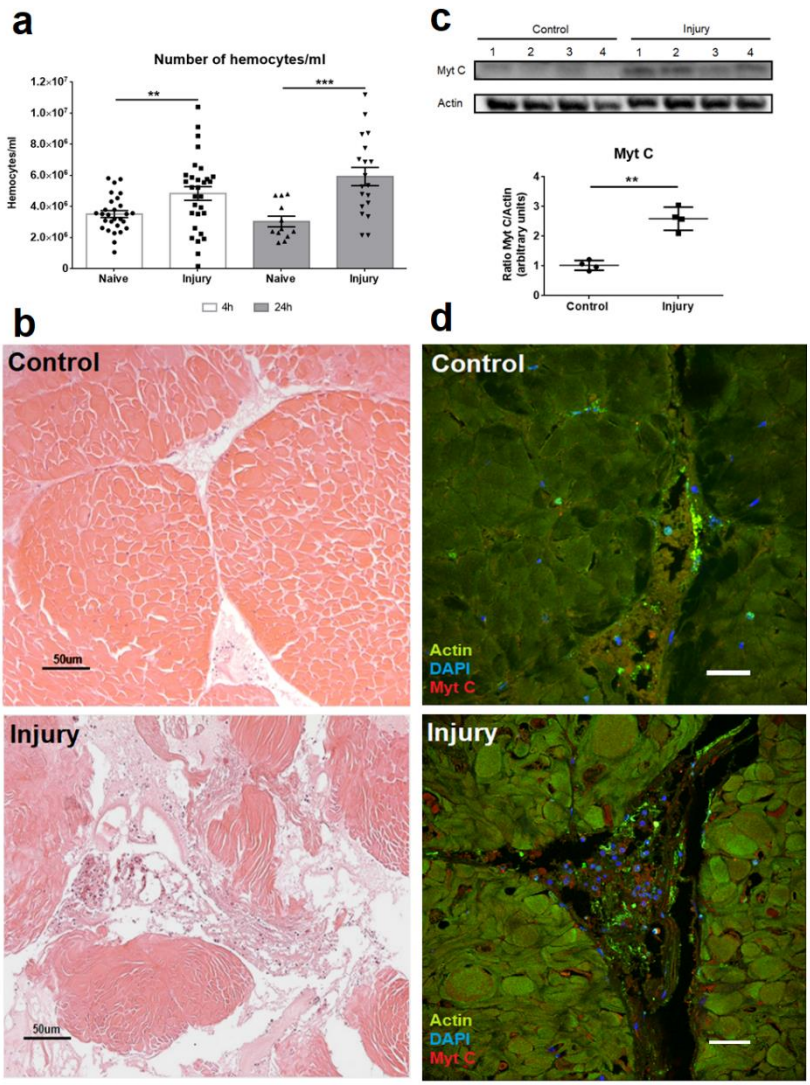

Figure 5. (a) Number of hemocytes in mussel muscle 4 h and 24 h after an injury. Asterisks indicate significant differences at $p < 0.01$ (**) and $p < 0.001$ (***). **(b)** Histological sections of healthy or injured adductor muscle using hematoxylin and eosin (HE) staining. Note the increase in the number of hemocytes in the damaged area. **(c)** Western blot of myticin C in samples of hemocytes extracted from naïve and injured mussels. Asterisks indicates significant differences at $p < 0.01$ (**). **(d)** Immunofluorescence of muscle sections of healthy or injured muscles. Triple labeling of actin (green), myticin C (red), and 4′,6-diamidino-2-phenylindole (DAPI) (blue). Bar scale represents 25 µm.

To determine if the migration of hemocytes would result in a greater presence of myticin C in the damaged area, a western blot was performed. Four pools of three mussels were injured in the adductor muscles, and the hemocytes were sampled. The result showed a significant increase of myticin C in hemocytes from mussels that had received an injury (Figure 5c). According to this, an immunofluorescence assay was performed using fixed adductor muscles, and it can be observed how the damaged area, again, was full of hemocytes, and how these cells exhibited higher levels of myticin C than controls (Figure 5d).

4.3.5 Effect of myticin C on human keratinocytes wound healing and zebrafish tail fin regeneration

The chemotactic properties of myticin C and its higher concentration around wounds suggest that it could be involved in accelerating the regeneration processes. With regard to this idea, a wound healing assay was performed with human HaCaT cells and zebrafish larvae, to investigate the myticin C potential in other species. Human keratinocytes showed a better ability to close the wound in a cell monolayer in presence of myticin C (Figure 6a). An accelerated regeneration from the first time assessed (24 h after the treatment) can be observed. However, a significant effect was observed 48 h after myticin C treatment (Figure 6b).

An *in vivo* test using wild-type zebrafish larvae was performed. The larvae tail fin was amputated, and after treatment with myticin C, the regeneration was followed for seven days. A significant increase of tail regeneration was observed after the second day, reaching the maximum at day 4 (Figure 6c). The measures were performed as Figure 6d illustrates.

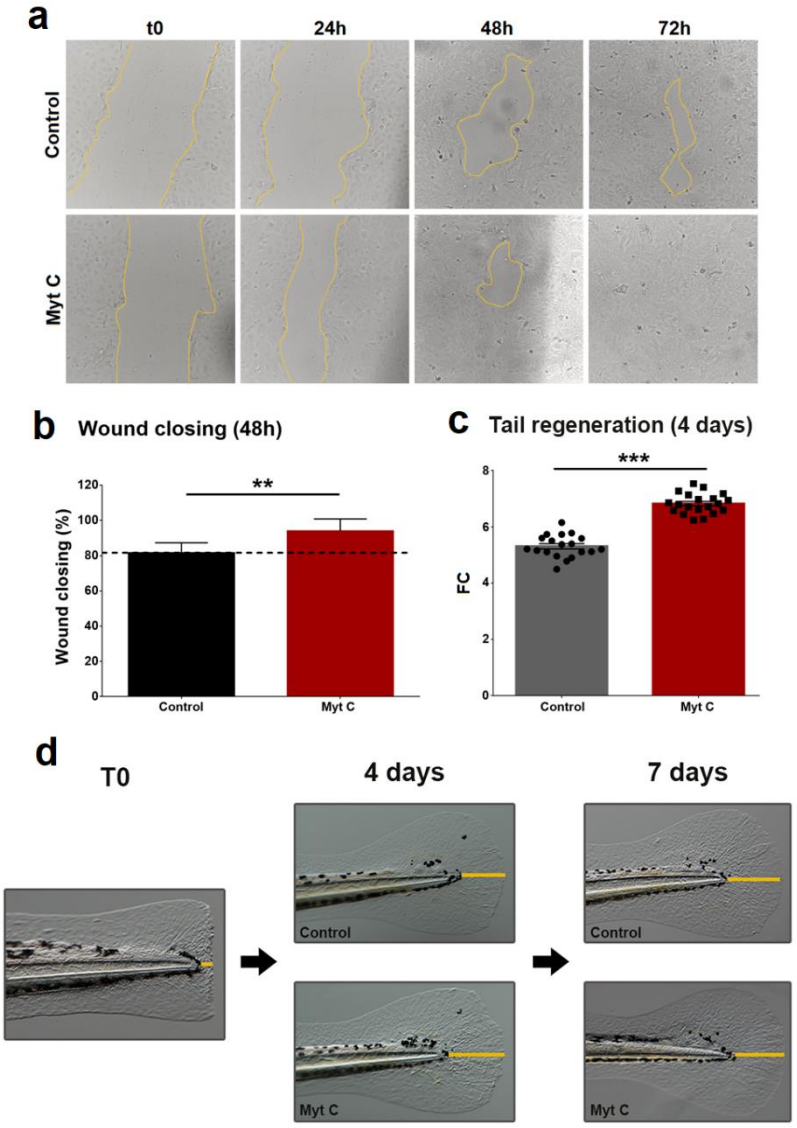

Figure 6. **(a)** Progress of the closure of the gap in HaCaT cells. **(b)** Percentage of wound closing 48 h after myticin C treatment. Asterisks indicate significant differences at $p < 0.01$ (**). **(c)** Tail regeneration, compared to time 0 (tail amputation), four days after the treatment. Asterisks indicate significant differences at $p < 0.001$ (***). **(d)** Images of a representative wild-type (WT) zebrafish larva after the tail fin amputation, and its regeneration after the treatment.

4.4 DISCUSSION

Myticin C is an antimicrobial peptide widely studied in mussels. Some of its functions are directly related to killing pathogens, being reported its antibacterial [2] and antiviral properties [4,15]. In line with these properties, one of the most expressed genes in the transcriptome analysis is the tripartite motif-containing protein 56 (TRIM56). The TRIM family of proteins is induced by interferon in vertebrates, and can limit viral growth [22].

Nevertheless, myticin C function is not limited to its antimicrobial activity. *In vitro* chemotactic properties of myticin C had been reported [15], and this capacity to promote cell migration makes myticin C a chemokine-like molecule. In our results, this function could also be observed through the up-regulation of dipetalogastin, a thrombin inhibitor well defined in the blood-sucking insect *Dipetalogaster maximus* [23,24]. This thrombin activity has been related to cancer and metastasis. The inhibition of these enzyme would be associated to increased cell metastatic behavior [25], which indicates that cellular movement is one of the regulated processes in myticin C-treated hemocytes. Cellular mobility implies conformational changes of the actin cytoskeleton [26]. Some proteins described as modulators of actin cytoskeleton were also up-regulated after myticin C treatment: transgelin-like protein and calponin-like protein. These proteins are responsible for gelling actin contributing to the formation of podosomes, the structures that enable cell movement and invasion [27]. Hemocytes treated with myticin C showed cytoplasmatic structures and prolongations towards the outside of the cell, which would be reminiscent of these podosomes. A change of actin conformation in histological preparations of adductor muscle of mussels treated with myticin C has been observed. This formation of structures similar to podosomes, as well as any cellular movement that requires the balance and regulation of the actin cytoskeleton, needs myosin as the motor protein to transform chemical energy into mechanical energy (through ATPase activity) [28]. All myosins are composed of one or two heavy chains and several light chains [28]. Transcriptome results showed up-regulation of several myosin components (two myosin heavy chains and one myosin regulatory light chain), showing a motor effect of myticin

C on hemocytes. In this regard, COX1 and Cytoglobin-1, the most up-regulated transcripts, support the fact that hemocytes need energy supplies for taxis after myticin C stimulation. These genes are related to the respiratory chain, and therefore to the increase of ATP production [29,30].

All these signs point towards an increase in the motility behavior of the hemocytes in the presence of myticin C. These results were further confirmed with time-lapse imaging of the hemocytes after being treated. These cells are capable of traveling a greater distance than control hemocytes, in agreement with the transcriptomic results. This ability of cells to respond to specific signals that guide their movement has been widely studied in neural crest migration [31], inflammatory disease [32,33], or morphogenesis and regeneration [34]. In this sense, cell migration is a key process that is involved in very different biological processes, such as embryonic development, immune response, or regeneration [35–38].

Cell migration is the first action that happens after an injury. The production of growth factors and chemokines attract new cells in charge of removing debris and stimulate angiogenesis, as well as extracellular matrix production [39]. In mussels, hemocytes would be attracted by chemokine-like molecules, such as myticin C, and would play their role in the damage resolution. At the same time, the increase of hemocytes numbers would enhance the production and release of more myticin C in the damage area. This hemocyte behavior agrees with the new function of myticin C, as defined in the chapter 2 of this thesis, where a link between the up-regulation of myticin C and the response to DAMPs and tissue injury was observed.

After cell recruitment and growth factor production, proliferation and maturation processes occur to reach the resolution of the damage [38–41]. Wound-healing assays are common to assess if a compound accelerates the regeneration process [42–44]. Myticin C showed the capacity to accelerate the closing of a gap in HaCaT cell monolayers, as well as to increase the speed of regeneration in wild-type zebrafish larvae after a tail fin amputation. Therefore, myticin C seems to enhance the regeneration process in mammal, fish, and bivalve cells.

In terms of expression values, transcriptomic results also showed the up-regulation of genes related to tissue regeneration, such as DC-STAMP. This gene regulates osteoclast differentiation, and in consequence, bone resorption in vertebrates [45]. The removal of damaged material before the deposition of new tissue is one of the initial steps in the regeneration cascade [39], and DC-STAMP, regulating the differentiation of both osteoclasts and osteoblasts, is a marker of bone turnover to assess remodeling and tissue healing [45]. Moreover, HSPG2, an extracellular matrix protein involved in preserving the integrity of extracellular matrices, is also up-regulated. This gene encodes the proteoglycan perlecan, whose function is directly related to controlling the cellular phenotype [46]. These genes control the damaged cells' removal, as well as regeneration processes and tissue structure maintenance.

4.5 Conclusions

Myticin C has such an effect on hemocytes that causes changes in their expression profile and mobility behavior. These changes are of consequence to the great number of genes directly related to the actin cytoskeleton, which are modulated by the peptide. In addition, in the situation of tissue injury, myticin C seems to accelerate all processes of regeneration. This would support the already proposed theory that myticin C is a cytokine-like protein exclusive to mussels.

4.6 References

1. FAO. FAO Home Page. Available online: http://www.fao.org/home/en/ (accessed on 6 December 2019).
2. Costa, M.M.; Prado-Alvarez, M.; Gestal, C.; Li, H.; Roch, P.; Novoa, B.; Figueras, A. Functional and molecular immune response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes against pathogen-associated molecular patterns and bacteria. Fish Shellfish Immunol. 26, 515–523 (2009). doi:10.1016/j.fsi.2009.02.001.

3. Mitta, G.; Vandenbulckeb, F.; Rocha, P. Original involvement of antimicrobial peptides in mussel innate immunity. FEBS Lett. 486, 185–190 (2000). doi:10.1016/s0014-5793(00)02192-x.

4. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral activity of myticin C peptide from mussel: An ancient defense against herpesviruses. J Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

5. Hubert, F.; Noel, T.; Roch, P. A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem. 240, 302–306 (1996). doi:10.1111/j.1432-1033.1996.0302h.x.

6. Mitta, G.; Hubert, F.; Noël, T.; Roch, P. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur J Biochem. 265, 71–78 (1999). doi:10.1046/j.1432-1327.1999.00654.x.

7. Mitta, G.; Vandenbulcke, F.; Hubert, F.; Salzet, M.; Roch, P. Involvement of mytilins in mussel antimicrobial defense. J Biol Chem. 275, 12954–12962 (2000). doi:10.1074/jbc.275.17.12954.

8. Sonthi, M.; Toubiana, M.; Pallavicini, A.; Venier, P.; Roch, P. Diversity of coding sequences and gene structures of the antifungal peptide mytimycin (MytM) from the Mediterranean mussel, *Mytilus galloprovincialis*. Mar Biotechnol. 13, 857–867 (2011). doi:10.1007/s10126-010-9345-4.

9. 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. 36, 390–399 (2012). doi:10.1016/j.dci.2011.08.003.

10. Liao, Z.; Wang, X.C.; Liu, H.H.; Fan, M.H.; Sun, J.J.; Shen, W. Molecular characterization of a novel antimicrobial peptide from *Mytilus coruscus*. Fish Shellfish Immunol. 34, 610–616 (2013). doi:10.1016/j.fsi.2014.09.019.

11. Qin, C.L.; Huang, W.; Zhou, S.Q.; Wang, X.C.; Liu, H.H.; Fan, M.H.; Wang, R.X.; Gao, P.; Liao, Z. Characterization of a novel antimicrobial peptide with chitin-biding domain from *Mytilus coruscus*. Fish Shellfish Immunol. 41, 362–370 (2014). doi:10.1016/j.fsi.2014.09.019.

12. Leoni, G.; De Poli, A.; Mardirossian, M.; Gambato, S.; Florian, F.; Venier, P.; Wilson, D.N.; Tossi, A.; Pallavicini, A.; Gerdol, M. Myticalins: A novel multigenic family of linear, cationic antimicrobial peptides from marine mussels (*Mytilus* spp.). Mar Drugs 15, 261 (2017). doi:10.3390/md15080261.

13. Pallavicini, A.; Costa, M.M.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

14. Costa, M.M.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

15. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: A chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS ONE 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

16. Martinez-Lopez, A.; Encinar, J.A.; Medina-Gali, R.M.; Balseiro, P.; Garcia-Valtanen, P.; Figueras, A.; Novoa, B.; Estepa, A. pH-dependent solution structure and activity of a reduced form of the host-defense peptide myticin C (Myt C) from the mussel *Mytilus galloprovincialis*. Mar Drugs. 11, 2328–2346 (2013). doi:10.3390/md11072328.

17. Qiagen. Quiagen Digital Insights Home Page. Available online: http://www.clcbio.com (accessed on 6 December 2019).

18. Robinson, M.D.; Smyth, G.K. Small-sample estimation of negative binomial dispersion, with applications to sage data. Biostatistics. 9, 321–332 (2008). doi:10.1093/biostatistics/kxm030.

19. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21, 3674–3676 (2005). doi:10.1093/bioinformatics/bti610.

20. Fisher, R.A. On the interpretation of $\chi 2$ from contingency tables, and the calculation of P. J. R. Stat Soc. 85, 87–94 (1922). http://dx.doi.org/10.2307/2340521.

21. Schneider, C. A.; Rasband, W. S; Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. Nat methods. 9, 671-675 (2012). doi:10.1038/nmeth.2089.

22. Ozato, K.; Shin, D.M.; Chang, T.H.; Morse, H.C., 3rd. TRIM family proteins and their emerging roles in innate immunity. Nat Rev Immunol. 8, 849–860 (2008). doi:10.1038/nri2413.

23. Mende, K.; Petoukhova, O.; Koulitchkova, V.; Schaub, G.A.; Lange, U.; Kaufmann, R.; Nowak, G. Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect Dipetalogaster maximus cDNA cloning, expression and characterization. Eur J Biochem. 266, 583–590 (1999). doi:10.1046/j.1432-1327.1999.00895.x.

24. Mende, K.; Lange, U.; Nowak, G. Three recombinant serine proteinase inhibitors expressed from the coding region of the thrombin inhibitor dipetalogastin. Insect Biochem Mol Biol. 34, 971–979 (2004). doi:10.1016/j.ibmb.2004.06.006.

25. Bazzi, Z.A.; Lanoue, D.; El-Youssef, M.; Romagnuolo, R.; Tubman, J.; Cavallo-Medved, D.; Porter, L.A.; Boffa, M.B. Activated thrombin-activatable fibrinolysis inhibitor (TAFIa) attenuates breast cancer cell metastatic behaviors through inhibition of plasminogen activation and extracellular proteolysis. BMC Cancer. 16, 328 (2016). doi:10.1186/s12885-016-2359-1.

26. Velle, K.B.; Fritz-Laylin, L.K. Diversity and evolution of actin-dependent phenotypes. Curr Opin Genet Dev. 58–59, 40–48 (2019). doi:10.1016/j.gde.2019.07.016.

27. Gimona, M.; Kaverina, I.; Resch, G.P.; Vignal, E.; Burgstaller, G. Calponin repeats regulate actin filament stability and formation of podosomes in smooth muscle cells. Mol Biol Cell. 14, 2482–2491 (2003). doi:10.1091/mbc.e02-11-0743.

28. Lodish, H.; Berk, A.; Zipursky, S.L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Myosin: The actin motor protein. In Molecular Cell Biology, 4th ed.; Freeman, W.H., Ed.; EEUU: New York, NY, USA; Volume 29, pp. 126–133 (2000).

29. Hankeln, T.; Wystub, S.; Laufs, T.; Schmidt, M.; Gerlach, F.; Saaler-Reinhardt, S.; Reuss, S.; Burmester, T. The cellular and subcellular localization of neuroglobin and cytoglobin—A clue to their function? IUBMB Life. 56, 671–679 (2004). doi:10.1080/15216540500037794.

30. Timón-Gómez, A.; Nývltová, E.; Abriata, L.A.; Vila, A.J.; Hosler, J.; Barrientos, A. Mitochondrial cytochrome c oxidase biogenesis: Recent developments. Semin Cell Dev Biol. 76, 163–178 (2018). doi:10.1016/j.semcdb.2017.08.055.

31. Shellard, A.; Mayor, R. Chemotaxis during neural crest migration. Semin Cell Dev Biol. 55, 111–118 (2016). doi:10.1016/j.semcdb.2016.01.031.

32. Reynolds, H.Y. Lung inflammation: Role of endogenous chemotactic factors in attracting polymorphonuclear granulocytes. Am Rev Respir Dis. 127, 16–25 (1983). doi:10.1164/arrd.1983.127.2P2.S16.

33. Rovin, B.H.; Phan, L.T. Chemotactic factors and renal inflammation. Am J Kidney Dis. 31, 1065–1084 (1998). doi:10.1053/ajkd.1998.v31.pm9631856.

34. Friedl, P.; Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. Nat Rev Mol Cell Biol. 10, 445–457 (2009). doi:10.1038/nrm2720.

35. Collins, C.; Nelson, W.J. Running with neighbors: Coordinating cell migration and cell-cell adhesion. Curr Opin Cell Biol. 36, 62–70 (2015). doi:10.1016/j.ceb.2015.07.004.

36. Barriga, E.H.; Mayor, R. Embryonic cell-cell adhesion: A key player in collective neural crest migration. Curr Top Dev Biol. 112, 301–323 (2015). doi:10.1016/bs.ctdb.2014.11.023.

37. Huse, M. Mechanical forces in the immune system. Nat Rev Immunol. 17, 679–690 (2017). doi:10.1038/nri.2017.74.

38. Bellavia, G.; Fasanaro, P.; Melchionna, R.; Capogrossi, M.C.; Napolitano, M. Transcriptional control of skin reepithelialization. J Dermatol Sci. 73, 3–9 (2014). doi:10.1016/j.jdermsci.2013.08.007.

39. Kasuya, A.; Tokura, Y. Attempts to accelerate wound healing. J Dermatol Sci. 76, 169–172 (2014). doi:10.1016/j.jdermsci.2014.11.001.

40. Gosain, A.; DiPietro, L.A. Aging and wound healing. World J Surg. 28, 321–326 (2004). doi:10.1007/s00268-003-7397-6

41. Kondo, T.; Ishida, Y. Molecular pathology of wound healing. Forensic Sci Int. 203, 93–98 (2010). doi:10.1016/j.forsciint.2010.07.004.

42. Rodriguez, L.G.; Wu, X.; Guan, J.L. Wound-Healing Assay. In: Guan JL. (eds) Cell Migration. Methods in Molecular Biology™, vol 294. Humana Press. (2005) https://doi.org/10.1385/1-59259-860-9:023.

43. Jonkman, J.E.; Cathcart, J.A.; Xu, F.; Bartolini, M.E.; Amon, J.E.; Stevens, K.M.; Colarusso, P. An introduction to the wound healing assay using live-cell microscopy. Cell Adh Migr. 8, 440–451 (2014). doi:10.4161/cam.36224.

44. Bilgi, A.; Zumrut Biber Muftuler, F.; Akman, L.; Ilker Medine, E.; Tonbaklar Bilgi, P.; Kozgus Guldu, O.; Goksun Gokulu, S.; Tekin, V.; Cosan Terek, M. *In vitro* determination of wound healing potential of axonge. Wounds. 29, 209–214 (2017).

45. Chiu, Y.H.; Ritchlin, C.T. DC-STAMP: A key regulator in osteoclast differentiation. J Cell Physiol. 231, 2402–2407 (2016). doi: 10.1002/jcp.25389.

46. Martinez, J.R.; Dhawan, A.; Farach-Carson, M.C. Modular proteoglycan perlecan/HSPG2: Mutations, phenotypes, and functions. Genes. 9, e556 (2018). doi:10.3390/genes9110556.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 5

COMPARATIVE GENOMICS REVEALS A SIGNIFICANT SEQUENCE VARIABILITY OF MYTICIN GENES IN *Mytilus galloprovincialis*

Rey-Campos, M.[1]; Novoa, B.[1]; Pallavicini, A.[2,3]; Gerdol, M.[2]; Figueras, A[1]. Comparative genomics reveals a significant sequence variability of myticin genes in *Mytilus galloprovincialis*. Biomolecules. 10, 943 (2020). doi:10.3390/biom10060943. (Open Access).

[1]Institute of Marine Research (IIM), National Research Council (CSIC), Eduardo Cabello, 6, 36208, Vigo, Spain. [2]Department of Life Sciences, University of Trieste, Trieste, Italy. [3]Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, Trieste, Italy.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 5: Comparative genomics reveals a significant sequence variability of myticin genes in *Mytilus galloprovincialis*

5.1 INTRODUCTION

Myticins are cysteine-rich antimicrobial peptides, discovered in hemocytes and plasma of *Mytilus galloprovincialis* in 1999 [1]. Since then, information about their gene and protein structure, sequence diversity, and function have been reported, mainly for the highly variable myticin C [2–6]. However, these studies have probably raised so far more questions about the function and evolution of myticins than they solved.

The effectiveness of mussel immunity primarily relies on hemocytes, which produce a significant variety of antimicrobial peptides (AMPs). Transcriptomic data revealed that hemocytes constitutively express very high levels of myticins [2,7] and other small cationic and cysteine-rich peptides such as defensins [8], mytilins [9], and mytimycins [10]. Altogether, these molecules constitute a complex repertoire of AMPs that is thought to have been developed by evolution as an effective defense system against pathogens present in the marine environment.

Like other mussel AMPs, myticin genes comprise 4 exons and 3 introns. The first exon is the smallest one and it only includes the non-coding 5'UTR region. The second, third, and fourth exons encode the signal peptide (about 20 amino acids), the mature peptide (about 40 amino acids), and the C-terminal region (about 40 amino acids),

respectively [4]. Like several other AMPs, myticins are therefore produced as inactive precursors, stored in hemocyte granules and activated by the proteolytic cleavage of the C-terminal region upon infection [6]. The very first studies carried out on these molecules allowed for the identification of different myticin variants, enabling a first categorization of these sequences among three main isoforms (A, B, and C) [1,2]. The three isoforms showed slightly different primary sequences and similar lengths, with the exception of a 4-amino-acid-long insertion in the C-terminal part of myticin C. However, further studies have added additional complexity to this picture with the use of massive sequencing [11] and the aforementioned classification now appears to be outdated.

Despite the variability of mussel myticins, some amino acids found in their mature peptide are highly conserved. In particular, eight highly conserved cysteine residues are engaged in four intramolecular disulfide bridges, which define the typical cysteine-stabilized alpha-beta (CSαβ) motif shared by myticins and other AMPs, such as defensins and mytilins [12].

From a functional point of view, myticin A and B have been defined as antibacterial and antifungal peptides [1]. On the other hand, myticin C, the isoform subjected so far to more intense studies, has been linked not only to antibacterial [13], but also to antiviral and cytokine-like functions [5,6]. The antiviral activity of myticin C was evidenced against fish viruses (VHSV and IPNV) [5], oyster herpesvirus (OsHV-1), and even human herpesvirus (HSV-1 and HSV-2) [6]. Moreover, the chemotactic activity defined in 2011 by Balseiro et al. [5] attributed myticin C a chemokine role, which was further supported by our recent work in which it was proven that these peptides are involved in tissue injury and regeneration processes in *M. galloprovincialis* [Chapters 2 and 4 of this thesis].

The main objective of this work was to shed some light on whether the remarkable intraspecific sequence variability of myticins derives from the complex genomic architecture of *M. galloprovincialis* or from RNA editing, exploiting the new information obtained from the mussel genome and the massive resequencing data of 16 additional individuals

[14]. The high variability of these AMPs and other immune effectors may represent a key factor in explaining the great evolutionary success of this species.

5.2 MATERIALS AND METHODS

5.2.1 Searching, screening, and identifying *M. galloprovincialis* myticins

The recently published reference mussel genome [14], and the resequenced genome assemblies of 14 additional different individual genomes from Galicia and Italy (8 males and 6 females) were screened for the presence of myticins gene variants. In this work, these genomic resources will be labeled as follows: LOLA (the reference genome), PURA, GALF1, GALF2, GALF3, GALM1, GALM2, GALM3, GALM6, GALM11 (from Galicia), ITAF1, ITAF2, ITAF3, ITAM1, ITAM2, and ITAM3 (from Italy). "F" and "M" indicate female and male mussels, respectively; note that LOLA and PURA are female individuals. Briefly, as detailed in the original paper [14], the mussel reference genome was assembled through a hybrid multi-step process, which included 2 x 101 bp paired-end, mate-pair, and fosmid-end Illumina reads, generated on a HiSeq2000 platform, as well as long PacBio reads, generated on a Sequel platform. Overall, the Illumina PE and PacBio sequencing outputs accounted for ~110 X and ~10 X coverage, respectively. The final assembly underwent multiple rounds of scaffolding, decontamination from exogenous contaminants, and removal of duplicated haplotype blocks to obtain a haploid reference assembly. The *de novo* assemblies of the resequenced individuals were obtained with the CLC Genomics Workbench 20.0.3 (Qiagen, Hilden, Germany) starting from 2 x 150 bp Illumina reads, generated with a HiSeq2500 platform, which accounted for ~30–35 X coverage.

All 16 genome assemblies were used to create BLAST databases. Previously described myticins [1,2], and specifically the third exon (mature peptide coding region) were used as a query to perform tBLASTn searches against each genome assembly, with an e-value threshold of 1e-5, using the CLC Genomics Workbench 20.0.3 (Qiagen,

Hilden, Germany). The resulting hits were manually checked to verify that they were indeed myticins (myticin database available in https://www.mdpi.com/2218-273X/10/6/943/s1). The reliability of the sequences obtained was verified by the visual inspection of read mapping data obtained with strict mapping thresholds (set with the CLC Genomics Workbench 20.0.3) to ensure the lack of sequencing or assembly errors.

The *in silico* translated exon 3 sequences of all obtained sequences were aligned using MUSCLE in the MEGA-X Software environment [15].

All the nucleotide sequences were clustered by similarity using the CD-HIT server [16], setting a sequence identity cut-off of 0.95. A consensus sequence was then established for each cluster, enabling further downstream analyses.

5.2.2 Phylogenetic analysis

The whole set of nucleotide sequences was taken into account to find the best suitable molecular model of evolution. jModelTest [17,18] was the software used for this purpose, and the choice of the best-fitting model, i.e., a Jukes–Cantor model [19], was performed based on the corrected Akaike information criterion.

A Bayesian inference analysis was run with a Markov chain Monte Carlo approach using the MrBayes v3.2.7 Software [20]. Two independent analysis with four chains each were run in parallel for 600,000 generations until the effective sample size parameter estimated for all the parameters of the model reached a value >200. The resulting phylogenetic tree was graphically represented using FigTree v1.4.4. [21]. By the same approach, a phylogenetic analysis was also performed using the established consensus sequences for each of the clusters of mussel myticins.

5.2.3 Isoelectric point

The N-terminal end of the mature peptides was predicted based on the detection of the signal peptide cleavage site with SignalP v3.0 [22]. Due to the unknown nature of the protease involved in the cleavage of the C-terminal region of myticin precursors, the putative C-terminal end of the mature peptide was identified based on the alignment with the known mature peptides previously described by other authors [1].

The isoelectric point (pI) and the charge of the mature peptide (at cytoplasmic pH = 7.4) of all the conventional myticins (myticins that display the usual 8-cysteine array in the mature peptide) was calculated using the Isoelectric Point Calculator Software [23]. Moreover, pI distribution of the complete peptide of the three classically defined myticins A, B, and C was analyzed through the calculation of the average pI based on a sliding window of 15 amino acids.

5.2.4 Positive and negative selection analysis

The codon-aligned nucleotide sequences of the exon 3 of the whole set of conventional myticins were analyzed to detect sites evolving under episodic positive selection with the MEME algorithm [24], as well as pervasive positive/negative selection using FEL [25], FUBAR [26], and SLAC [25] algorithms. A similar analysis was also performed on a subset of sequences belonging to the myticin C clade only. These analyses were performed using Datamonkey Adaptive Evolution Server [27]. The predicted three-dimensional structure of mature peptide of myticin C was obtained from a previous publication [28] and modified with Chimera 1.14 [29] to highlight sites under significant positive and negative selection (i.e., p-value lower than the 0.1 threshold).

5.2.5 Promoter analysis

The previously identified exon 3 matches were used as a seed to extend the reconstruction of full myticin genes from the genome assemblies, retrieving the sequences of exons 1, 2, and 4, whenever possible. The Genie tool [30] was used to predict the 5’ splicing

acceptor sites and the 3' splicing donor sites, and thereby to define the boundaries of cited exons, whenever it was possible. Once the full gene structure was appropriately annotated (15 out of 32 clusters, full sequences in https://www.mdpi.com/2218-273X/10/6/943/s1), we extracted a fragment of 500 bp upstream of the first exon to perform a promoter analysis by searching for conserved ungapped motifs shared by most myticin genes. This length threshold was selected as a compromise between the inclusion of a significant number of sequences in the analysis and the possibility to explore a biologically meaningful sequence context (i.e., the core and proximal promoter), based on the suggestions provided by Zia and Moses [31] to limit false positive detection. Obviously, the weakness of this approach, also linked with the relatively high fragmentation rate of the mussel genome reassemblies, was the impossibility of investigating the presence of distal regulating elements present upstream or downstream of the transcription start site. The *de novo* motif-finding analysis were run with MEME Suite 5.1.1 [32], selecting the classic motif discovery mode and setting the accepted length of such motifs between 6 and 30 bp.

The significant motifs obtained (combined match p-value lower than 1e-15) were kept for a further search in the LOLA assembly (the reference genome assembly, [14]) in order to determine if the motifs identified in the myticin promoter region were also associated with other mussel genes. For this, the motif search tool of CLC Genomics Workbench 20.0.3 was used. A list containing the 10 consensus motif sequences of myticin was created and the subsequent search was performed with 70% accuracy.

5.2.6 Genomic vs. transcriptomic data

In addition to the genome of LOLA, its transcriptome was also available [14]. LOLA myticin sequences found in the genome were compared to all sequences present in the transcriptome (the approach to search the sequences in the transcriptome assembly was the same as the genomic approach, previously described). Genome reads were mapped to the myticin sequences found in the transcriptome in order to determine differences between the DNA and RNA of the same

individual. The mapping parameters were set to be highly restrictive in order to only allow perfect matches (length fraction = 1 and similarity fraction = 1). The mapping files obtained were then visually inspected to detect regions with no read coverage, which could indicate mismatches between the genomic DNA and mRNA sequences and pinpoint the presence of sites subjected to RNA editing.

5.2.7 Expression analysis

Taking advantage of the transcriptomic information available in the Sequence Read Archive, National Center for Biotechnology Information (SRA-NCBI), 6 different transcriptome assemblies of *M. galloprovincialis* (from different geographic locations and different tissues) were used to find evidence of expression of all the different clusters of myticin. These transcriptomes are the following: PRJNA88481 (digestive gland, which also included unpublished gill data available at the University of Trieste), PRJNA525609 (mantle), PRJNA249058 (whole body), PRJNA484309 (gills and mantle), PRJNA230138 (hemocytes, mantle, muscle, and gills) and PRJNA466718 (hemocytes). These resources were screened with the tBLASTn approach, using the mature peptide regions of the aforementioned myticin clusters as a query, as previously described. In this case, the finding of a match sharing >95% identity with the query sequence was considered as an evidence of the expression of a myticin variant belonging to the underlying sequence cluster. All the additional contigs identified that displayed >5% divergence compared with the established clusters were considered as belonging to new unreported clusters and added to the myticin sequence dataset (available in https://www.mdpi.com/2218-273X/10/6/943/s1).

5.3 RESULTS

5.3.1 Searching, screening, and identifying *M. galloprovincialis* myticins

A total of 120 different nucleotide sequences encoding myticins were found in the 16 mussel genome assemblies. From these 120 sequences, 93 were conventional myticins (myticins that display the usual 8-cysteine array in the mature peptide), 9 were pseudogenes (sequences that incorporate a STOP codon which interrupts the open reading frame), and 18 were pseudomyticins (sequences that keep the most of the structure of myticins but lose a pair of cysteines) (all the sequences are available in https://www.mdpi.com/2218-273X/10/6/943/s1). All the sequences were clustered based on an identity percentage threshold of 95%, obtaining a total of 32 different clusters (clusters of sequences are available in https://www.mdpi.com/2218-273X/10/6/943/s1).

5.3.2 Phylogenetic analysis

The Bayesian tree of all the 120 myticin variants identified in this study (Figure 1) displayed a remarkable sequence diversification, with a subdivision of classical myticins (A, B, and C) and pseudomyticins, as well as other newly reported myticins belonging to intermediate branches.

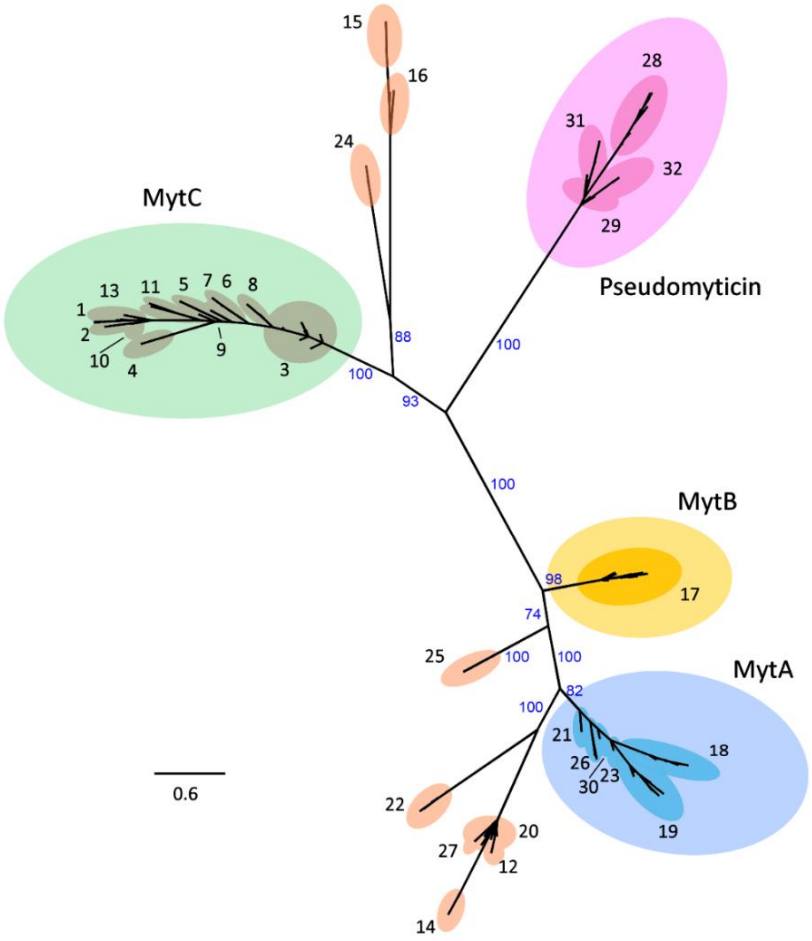

Figure 1. Phylogenetic analysis. Exon 3 of 120 myticin sequences obtained from 16 genomes were subjected to multiple sequence alignment and analyzed with Bayesian phylogenetic inference. The employed evolution model was the Jukes and Cantor model (JC). Black numbers at termini correspond to the 32 major groups (based on a pairwise sequence identity >95%). Green (MytC), yellow (MytB), blue (MytA), and purple (pseudomyticin) ellipses indicate the variants falling within the four main described groups of myticins and pseudomyticins. Blue numbers correspond to the posterior probabilities of the branches.

To facilitate the analysis, a simplified version of myticin phylogeny, showing the evolutionary relationship among myticin clusters, is displayed in figure 2a.

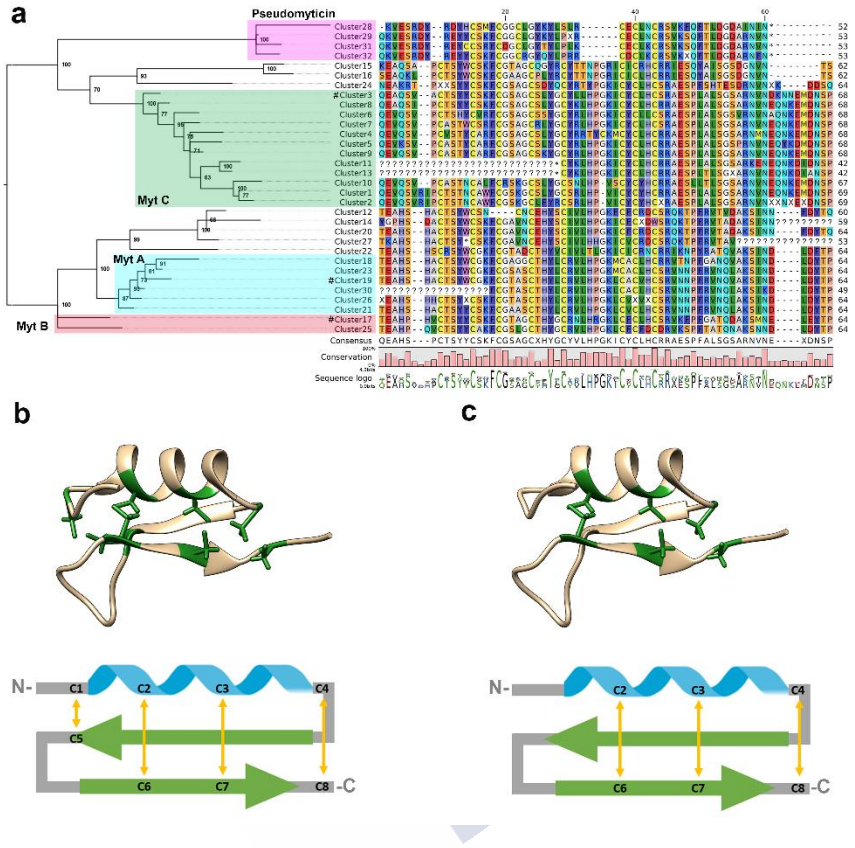

Figure 2. Clusters phylogenetic analysis and alignment. **(a)** A consensus sequence of each of the 32 myticin clusters (exon 3) was used to run the phylogenetic analysis and build the alignment. ? represents parts of sequences that likely resulted from an exon truncation event. X represents ambiguous sites in the consensus. * represents STOP codons. Green (MytC), red (MytB), blue (MytA), and purple (pseudomyticin) squares show the position of the four main described groups of myticins and pseudomyticins. # represents previously published Myt A/B/C sequences. **(b)** and **(c)** show the predicted secondary structure (green color highlights the cysteine positions) and the putative interaction of cysteines engaged in the formation of the disulfide bonds (yellow arrows) of conventional myticins and pseudomyticins, respectively.

The same figure also displays an alignment of the consensus sequence of the 32 myticin clusters, which allows us to note the greatest differences between pseudomyticins and the rest of myticins. This diversified group has lost 2 out of 8 characteristic cysteines of this gene family, namely Cys1 and Cys5, which are expected to be engaged in one of the four disulfide bonds of the CSαβ structural scaffold of myticins. Two additional panels of figure 2 show the expected disulfide array of conventional myticins (Figure 2b) and pseudomyticins (Figure 2c).

5.3.3 Presence/absence variation

An evaluation of the presence/absence of all 120 sequences was performed in the 16 mussel genomes (presence/absence matrix is available in https://www.mdpi.com/2218-273X/10/6/943/s1). On average, each mussel genome showed 11 myticin different sequences, of which around 6 were exclusively found in one out of the 16 individuals analyzed. The presence/absence matrix highlights the great inter-individual diversity in the repertoire of myticins of each individual, which results in a virtually unique collection of variants in each mussel. The sequences located at the top of the matrix are those which displayed the highest frequency of occurrence (2, 98_PM, 1, and 18). Even so, none of them was present in all the genomes analyzed.

As several of the 120 variants identified only displayed minor differences in pairwise comparisons, we cannot exclude that they represent polymorphic alleles of the same gene. Our clusterization approach allowed us to take into account these uncertainties in the verification of presence/absence variation, identifying several groups of myticins shared by most of the 16 analyzed genomes and others which were just found in a low number of individuals, or were even exclusively present in a single one (Figure 3). It could be observed as at least one representative of each of the four major groups of myticins (A, B, C, and pseudomyticins) was present in all or almost all the genomes (clusters 8, 19, 25, and 28). Of all sequence clusters, only cluster 27 was present in all individuals.

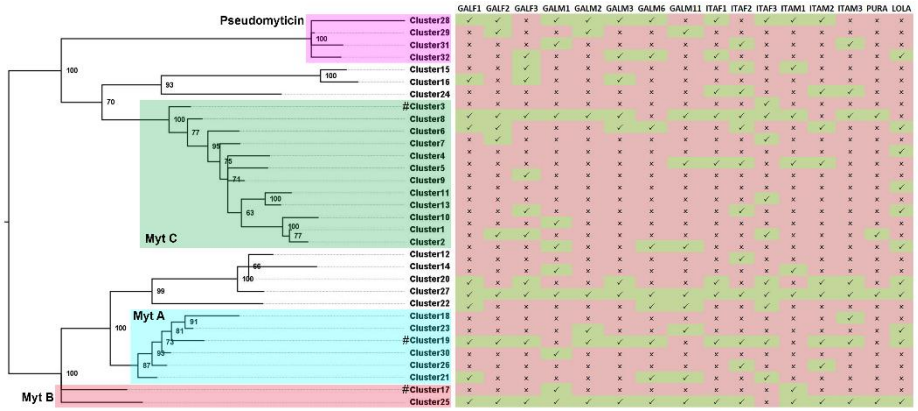

Figure 3. Presence/absence evaluation of the 32 clusters of myticin. The matrix shows presence/absence of each cluster in the 16 mussel genomes. Green (MytC), red (MytB), blue (MytA), and purple (pseudomyticin) squares show the position of the four main described groups of myticins and pseudomyticins. # represents previously published Myt A/B/C sequences.

5.3.4 Isoelectric point

The isoelectric point and predicted charge at cytoplasmic pH (i.e., 7.4) of all the conventional myticin sequences (mature peptide) was calculated. Figure 4a shows that all myticins display very narrow variations in terms of pI, which varies between 7 and 8. It can also be observed that neither pI, nor the molecular weight of the mature peptide depend on the myticin isoform. The charge of the mature peptides is usually slightly positive, varying from -2 to 4 (Figure 4b), with no remarkable differences between different myticin isoforms.

The sliding-window analysis of pI along the whole peptide, carried out on the three representative precursor peptides of myticin A, B, and C (Figure 4c), showed very similar profiles, with a stable value across the mature peptide, and just a slight decrease in the C-terminal part of the sequence, of a much smaller entity than previously observed in other mussel AMPs, such as mytilins [33].

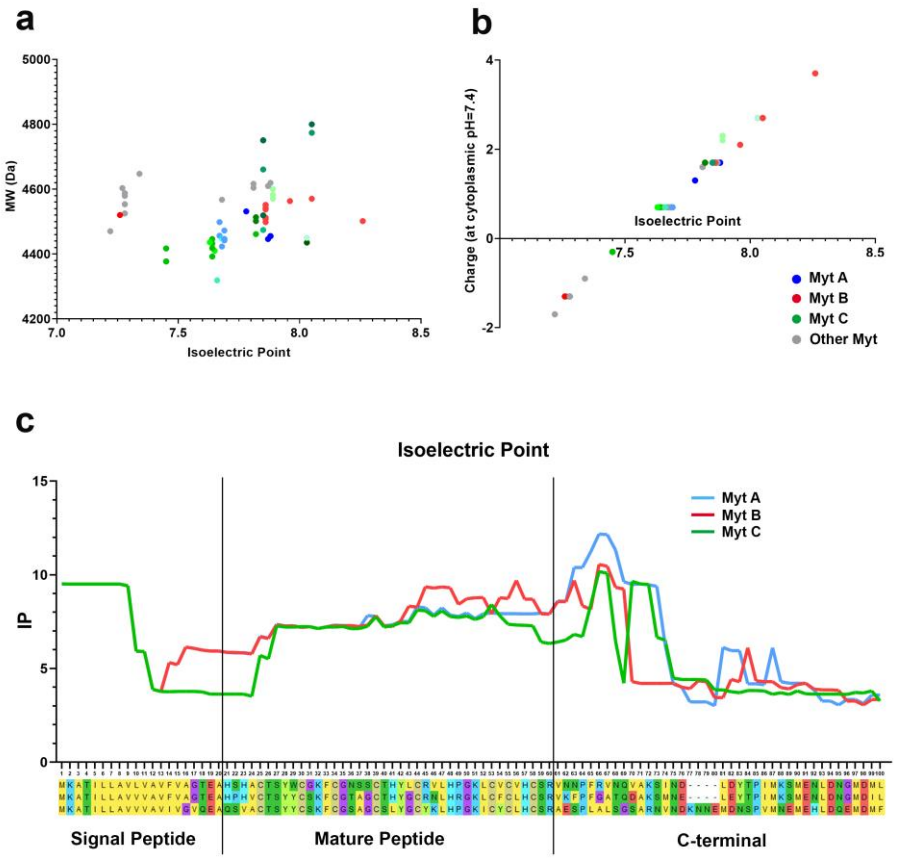

Figure 4. Isoelectric Point. **(a)** Isoelectric point (X axis) and molecular weight (Y axis) of the mature peptide of each conventional myticins (71 defined mature peptides, obtained by removing redundant amino acid sequences and pseudogenes). **(b)** Isoelectric point (X axis) and charge at pH = 7.4 (Y axis) of the mature peptide of each conventional myticins (71 defined mature peptide). **(c)** Isoelectric point of the whole sequence of one representative of each conventional myticin (A, B, and C). The isoelectric point distribution was analyzed through the calculation of the average isoelectric point based on a sliding window of 15 amino acids.

5.3.5 Positive and negative selection analysis

The selection analysis identified several sites subject to positive or negative selection in the mature peptide region. Despite some minor

differences, the various tests were concordant in recognizing multiple sites under pervasive purifying selection (Figure 5). These sites match with six of the cysteines that form the characteristic disulfide array of myticins. Specifically, Cys-1, -4, -5, -6, -7, and -8 are under a strong negative selection. Another interesting case of negative selection is the arginine that marks the end of the mature peptide, which may be recognized as the signal for proteolytic cleavage of the precursor protein. Three additional negatively selected sites of interest that emerged from this analysis are a conserved glycine found in a tight turn which connects the two antiparallel beta sheets, and a serine and a phenylalanine residue found in the alpha helical region. The tests also identified several sites evolving under significant pervasive positive selection (4 by FEL, 5 by FUBAR, and 2 by SLAC). MEME further indicated the possibility that 10 out of the 42 amino acids in the myticin mature peptide might have undergone episodic positive selection. The functional and structural role of these hypervariable sites is presently unknown. The same analysis, run on the sequences belonging to the highly variable myticin C clade only, revealed a good overlap of selected sites compared with the full sequence dataset, supporting the reliability of the results described above and pointing out that the signals obtained were not just the result of ancestral divergence among paralogs.

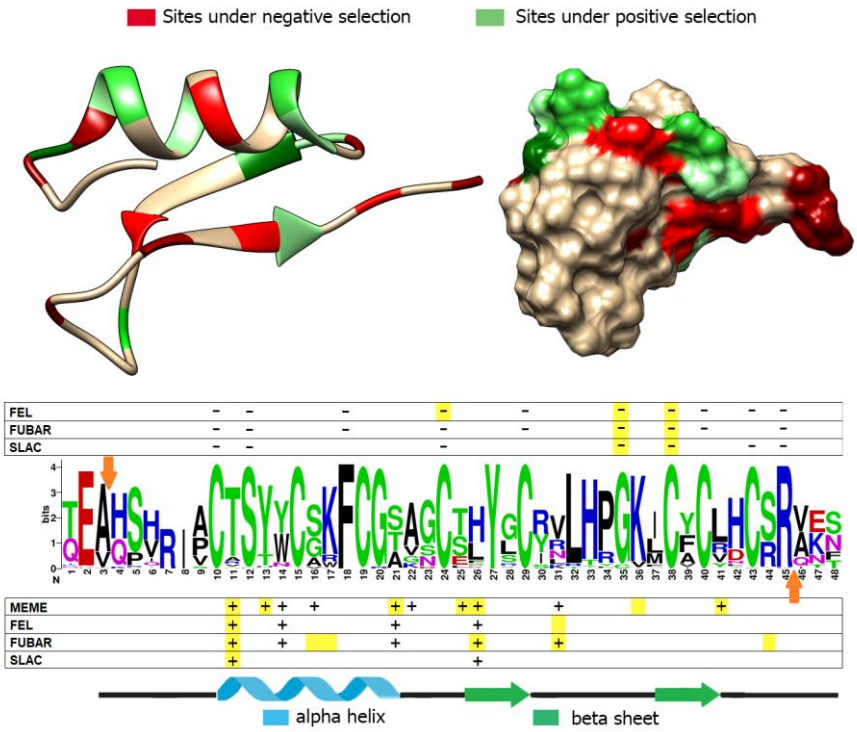

Figure 5. Positive and negative selection analysis. Exon 3 of all conventional myticins were selected to perform the analysis, discarding pseudogenic sequences. Four prediction models have been used (MEME, FEL, FUBAR, and SLAC). Positively selected sites (+) and negatively selection sites (-) identified are both shown on the graph. Positively and negatively selected sites identified in the myticin C sequence clade are marked with a yellow background. Orange arrows point to the beginning and end sites of the mature peptide. The ribbon structure and the molecular surface, highlighting in red the sites under negative selection and in green the sites under positive selection, are also shown [from 28].

5.3.6 Promoter analysis

The promoter analysis allowed to find 10 different motifs located in the 500 bp upstream of the gene transcription start site, which was expected to include the core and proximal promoter elements (Figure 6). These serial motifs were found in most myticin genes and displayed a nearly invariable position. This observation, together with the high primary sequence divergence among variants, suggests the presence of

a well-defined promoter architecture shared by all myticins. Specifically, motifs 2 (CAACCACAATKTCCGTSTTTCCTGTYWAGA) and 4 (AAAARWMGDMTAYTACGCARAWAKATTTKG) were found in all the tested sequences. On the other hand, motifs 1 (ATATATAYATWATAYTWATATACATKTCT) and 3 (CMAAAWACTAYGCTTTTAAATMTAATGCAG) appeared to be associated with each other and were only found in myticin C and evolutionarily-related sequences. Other least conserved motifs were identified in a lower number of sequences, but always displayed high positional conservation.

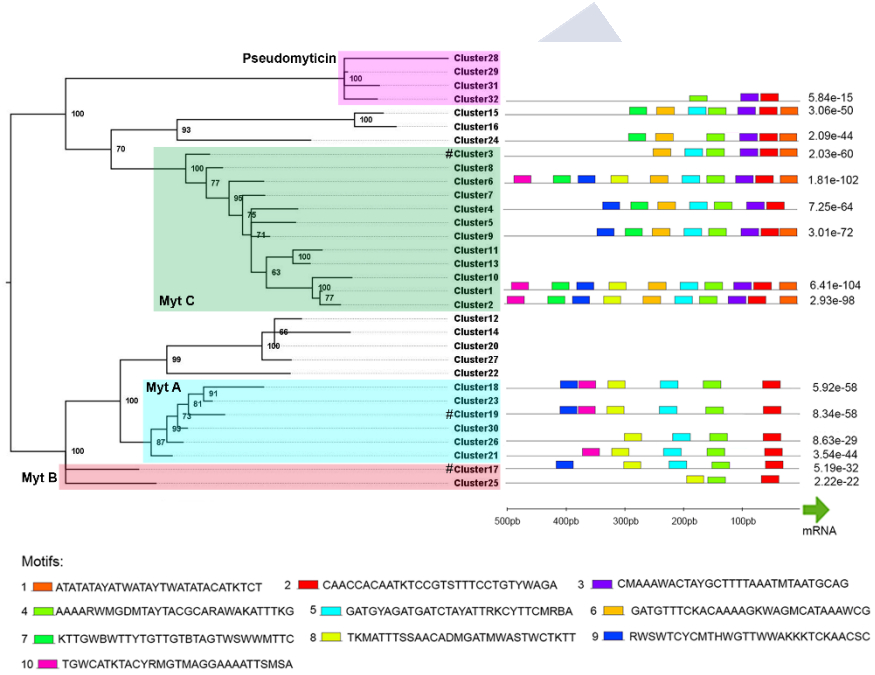

Figure 6. Promoter analysis. The 500 bp-long region immediately upstream of the transcription start site, and corresponding to the putative promoter of 15 complete available clusters, was analyzed with MEME. Colored boxes show the 10 different significant motifs found (the sequence of each motif is available in the legend). The p-value derived from the combined observation of all the motifs present in each sequence is also shown. The green arrow marks the beginning of the mRNA sequence. Green (MytC), red (MytB), blue (MytA), and purple (pseudomyticin) squares show the position of the four main described groups of myticins and pseudomyticins. # represents previously published Myt A/B/C sequences.

Although information about the transcription factor binding sites of any molluscan species is currently unavailable in specialized repositories, the strong hemocyte-specificity and high transcriptional activity of myticin genes in physiological conditions [2,7] most certainly suggests that their gene expression is strictly regulated by highly specific and likely unknown transcription factors, which may recognize some of the motifs described above. While we believe our observations may represent genuine transcription factor binding sites candidates, this *in silico* analysis should be complemented in the future by functional validation, i.e., by the identification of the transcription factors responsible for the regulation of myticin gene expression.

Unfortunately, due to the technical limitations linked with the fragmented nature of the genome assemblies, it was not possible to investigate whether any additional distal regulatory element was associated with myticin genes, either upstream or downstream of the transcription start site.

However, the conserved nature of the myticin promoter prompted us to investigate whether the same motifs could be identified in other genomic regions, associated with the promoter of other AMP gene families (e.g., defensins and mytilins), or with other genes with strong hemocyte-specific expression. However, the screening of the mussel reference genome did not reveal any other gene associated with the 10 aforementioned conserved sequence motifs.

5.3.7 Genomic vs. transcriptomic data

Having the genome and transcriptome of the same individual offers a great opportunity to investigate RNA editing processes occurring after the transcription [14]. Although eleven different myticin sequence variants were found in the reference genome of LOLA (presence/absence matrix available in https://www.mdpi.com/2218-273X/10/6/943/s1), only 7 different contigs were present in the transcriptome assembly obtained from the same individual, indicating that four of these were not expressed. The comparison between the genomic and RNA sequences of the seven expressed myticin variants

highlighted that there were no discrepancies. This result rules out the possibility that the high level of intraspecific sequence diversity of myticins derives from the mRNA editing process, confirming its genomic origins.

5.3.8 Expression analysis

A total of 6 different transcriptomes deriving from different mussel tissues and geographical locations have been analyzed to investigate whether any of the 32 previously defined sequence clusters were broadly expressed (Figure 7). Again, at least one cluster belonging to the four main groups of myticins (A, B, C, and pseudomyticins) was expressed in most of the analyzed transcriptomes (in particular, sequences from the cluster 3, 17, 19, and 32 were expressed in almost all the analyzed transcriptomes). Some other clusters (most of the cases) were scarcely expressed or even not expressed at all in any of the studied transcriptomes. For instance, clusters 1, 2, 8, 9, 14, 18, 23, and 25 were expressed in only one out of 6 transcriptomes. Some of the clusters lacking evidence of expression are, most likely, pseudogenes, as evidenced by the truncation of the open reading frame due to frameshift or nonsense mutations (Figure 2).

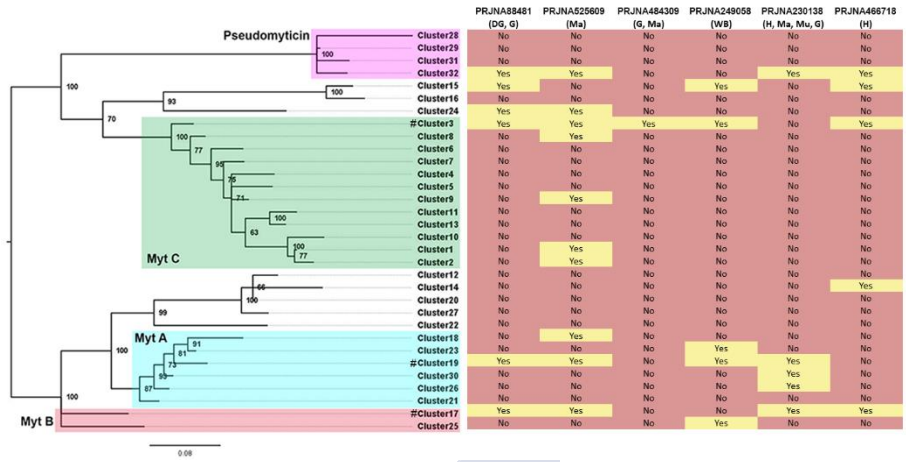

Figure 7. Expression analysis. A total of 6 transcriptomes derived from different tissues and mussels sampled in different geographical locations were analyzed to verify the expression of variants belonging to each cluster. The cluster expressed in each transcriptome is highlighted with a yellow background. Red color indicates no evidence of expression. Abbreviations mean: DG, digestive gland; G, gills; Ma, mantle; WB, whole body; H, hemocytes; Mu, muscle. Green (MytC), red (MytB), blue (MytA), and purple (pseudomyticin) squares show the position of the four main described groups of myticins and pseudomyticins. # represents previously published Myt A/B/C sequences.

Several of the sequences identified in these transcriptomes showed >5% primary sequence divergence compared with the clusters identified in this study. This observation indicates that the 16 mussels we analyzed, belonging to just two different populations, were not sufficient to build a complete collection of all the possible sequence variants found in the different *M. galloprovincialis* populations across the globe. This suggests that the sequence collection presented in this work might need to be updated in the future with the addition of novel variants and clusters.

5.4 Discussion

Myticins have been traditionally classified in three main groups, myticin A, B, and C. This classification attended to their amino acid sequence and function [1,2]. However, important technological

advantages related to massive sequencing techniques have generated additional information that is progressively revealing that the molecular diversity of myticins is much larger than previously thought [14].

The analysis of 16 fully resequenced mussel genomes allowed to ascertain that, on average, each animal possesses 11 different myticin variants, and that a large number of such isoforms are found at very low frequencies in mussel populations. This finding provides strong evidence in support of the hypothesis that each mussel contains its own unique repertoire of myticins, as previously suggested by Costa et al. [3]. These observations are consistent with gene presence/absence variation (PAV), which has been extensively studied in prokaryotes [34] and some plants [35], but only occasionally reported in metazoans. The Mediterranean mussel is the first metazoan where PAV has been described as a widespread phenomenon. Indeed, nearly one-third of the mussel protein-coding genes are subjected to PAV, meaning that they can be either present or absent in different individuals [14]. This study points out that the myticin gene family is also strongly affected by PAV, which appears to be the most important source of intraspecific genetic variation in this case [36]. It is important to note that our phylogenetic analysis revealed that the sequence variability of myticins covered a broad and nearly-continuous spectrum of diversity, preventing precise discrimination of each variant between allelic variants of the same gene and paralogous gene copies. Our clustering approach, based on an arbitrary 95% pairwise identity threshold, therefore needs to be considered with caution, since some of the 32 clusters identified may represent groups of divergent allelic variants. Nevertheless, the high number of variants identified in each mussel, their significant primary sequence divergence, as well as the multi-gene architecture of the myticin gene locus in the reference assembly [14], most certainly indicate that the myticin gene family comprises multiple paralogous genes.

Among the most significant findings of this study, we can report the presence of a new group of peculiar sequences, named pseudomyticins. These encoded peptides are characterized by the loss of the first and fifth cysteines of the typical disulfide array of myticins, which would result in the retention of just three out of the four disulfide

bridges described in classical myticins [12]. Despite these unusual features, pseudomyticins appear to be potentially functional genes, as evidenced by the maintenance of conserved motifs in the promoter region, as well as evidenced by their translation to mRNA collected from transcriptome data.

On the other hand, several other myticin variants, which lacked evidence of expression, were characterized by the truncation of the open reading frame, either due to the presence of frameshift/nonsense mutations or due to exon loss. This observation may be consistent with the progressive loss-of-function of some accessory myticin variants generated by past gene duplication events, the fate of which might have headed towards pseudogenization. This phenomenon has been previously observed in mussel mytilins, myticalins, big defensins, and mytimycins [14], as well as in other AMPs from diverse animals whose evolutionary diversification has been driven by gene duplication [37,38].

Like defensins [8,39] and mytilins [33,40], myticins belong to the CSαβ peptide superfamily, which includes several other structurally convergent AMPs found in other domains of life. The conservation of at least three out of the four disulfide bonds in the cysteine array of the mature peptide is essential for the maintenance of the CSαβ structural scaffold. Moreover, most CSαβ peptides have a cationic and amphipathic nature, which is thought to facilitate their electrostatic interaction with the negatively charged surfaces of gram-negative (outer membrane) or gram-positive (cell wall) bacteria [41,42]. The calculated isoelectric point (pI) and net charge at the physiological pH of all myticins were therefore surprisingly low, considering their hypothesized function as AMPs. Compared for instance to mytilins [33], that show an isoelectric point between 9–12 and an average net charge of +9, myticins just reached a maximum pI value of 8 and, in most cases, they only had a slightly cationic nature (with a predicted charge ranging from -2 to +4 in cytosolic conditions).

Similar considerations can be extended to the whole sequence of the precursor peptides. Classically, in several AMPs the signal peptide region is neutrally charged, while the mature peptide region displays a

strong positive charge, counterbalanced by the negative charge of the C-terminal region [33,42]. However, unlike mytilins and defensins, the pI profile of myticins was quite stable and only showed a significant drop in the charge in the final part of the C-terminal region. These charge distribution properties, unusual for an AMP, would find a justification in the reports that have recently suggested that myticins may cover additional functions, besides pathogen killing. The first indication pointing towards this direction came from the study of Balseiro et al. [5], which proposed myticin C as the first chemokine-like molecule in mussels, but new evidence now supports the chemotactic activity of this molecule. In fact, the expression of myticin was found to significantly increase in mussel after tissue injury, an effect which was not observed in the presence of a pathogen (*Vibrio splendidus*) [Chapter 2 of this thesis]. Moreover, a correlation between the expression of myticin after a tissue injury and the number of hemocytes recruited at the damaged area was also demonstrated. These observations allowed us to formulate a new functional hypothesis for myticin as a driver of tissue regeneration, as it is described in the chapter 4 of the present thesis.

The functional importance of cysteines is reflected from an evolutionary point of view. Despite the great variability and complexity of the mussel genome [14] and the enormous intrinsic variability of the myticins, the cysteine array of the mature peptide remained unchanged (with the aforementioned exception of pseudomyticins). Our selection analysis confirmed that this remarkable conservation derives from strong purifying selection. In addition, we provide evidence in support of the strong impact of purifying selection on four additional sites: Ser and Thr residue parts of the alpha helix region, a Gly included in the tight turn connecting the two antiparallel beta sheets and an Arg that limits the C-terminal boundary of mature peptide, which we hypothesize might serve as the site for the proteolytic cleavage of the precursor. In line with a previous report from Padhi and Verghese [43], our analysis revealed the presence of a significant number of sites evolving under diversifying selection. Our approach indicates that up to 25% of the sites included in the mature peptide region are subject to positive selection. This observation, supported by the significant

overlap with the sites detected with a parallel analysis carried out on myticin C variants only, indicates that positively selected sites are the key sites responsible for the high levels of myticin intraspecific diversity.

Although several studies have previously analyzed the molecular diversity of myticins [2–4,11,43], most of these suffered from important limitations, which had so far not been permitted to disclose the basis of these observations. These include: (i) The analysis of cDNA sequences only, which prevented the observation of non-expressed or poorly expressed variants; (ii), the use of PCR amplification, which might have introduced biases with primer design; (iii) the frequent use of data derived from pools of different individuals, which prevented any reliable assignment of variants to individuals; (iv) the lack of paired genomic DNA and mRNA sequence data.

The experimental design of this study avoided all the aforementioned issues and the comparison between the genomic DNA and mRNA sequences obtained from the same individual enabled us to establish that the huge level of sequence variability of myticins has an entirely genomic origin. RNA editing, i.e., the process of post-transcriptional modification of mRNAs through the inclusion of indels or the substitution of nucleotides, common in other molluscs such as cephalopods [44], does not seem to play any role in the generation of sequence diversity in myticins.

In terms of expression, myticins are highly expressed in different developmental stages [45] as well as in mussel hemocytes, where they emerge in the top 10 most actively transcribed genes [2,7]. These observations suggest that a strong core of regulatory elements, including promoters and enhancers, would regulate the expression of myticin genes. The way in which the transcription factors recognize these regulatory elements is far from fully understood [46,47]. In eukaryotic genomes, thousands of genes that encode messenger RNA are transcribed by the RNA polymerase II (POL II) molecular machinery. To initiate the transcription process, RNA polymerase recognizes the promoter region, located immediately upstream of the transcription start site of each gene. Some general motifs recognized by

POL II are the B recognition element (BRE), TATA box, initiator (Inr), motif ten element (MTE), and downstream promoter element (DPE) [48]. Although the TATA box is one of the most studied motifs in vertebrates [49,50], other CpG motifs represent other common promoter elements found in vertebrate genomes [51]. In general, TATA boxes tend to be associated with focused transcriptional initiation, whereas CpG motifs tend to display dispersed initiation patterns [52,53]. Even though many eukaryotic core promoters contain some of these motifs, no universal motif has ever been identified as unambiguously present in a core promoter in a given eukaryotic genome [54]. Moreover, to the best of our knowledge, no comprehensive study has ever been carried out to characterize the transcription factor binding sites found in Mollusca or, more broadly, in the Lophotrochozoa superphylum.

Our analysis of the 500-bp sequence upstream of the TSS of myticins allowed to identify 10 different conserved motifs that may be involved in the regulation of myticin gene expression as core and proximal promoter elements. These motifs were found in a variable number, but in a well conserved order, in canonical myticin genes, as well as in pseudomyticins. Although a putative TATA box could be recognized among these motifs (i.e., motif 1), as described above, no universal promoters have been determined yet. The conserved motifs defined in the myticin promoter are apparently not shared with other mussel genes, including other hemocyte-specific AMPs, like mytilins or defensins, which suggests that the expression of myticins may be controlled by highly specific and still uncharacterized transcription factors. Unfortunately, the limitations posed by the fragmented nature of the individual mussel genome assemblies prevented a detailed characterization of the distal regulatory elements that may contribute to this transcriptional regulation. Taking into account the scarcity of data available about transcription factor binding sites in Lophotrochozoa, the identification of the 10 conserved motifs reported above might provide a solid basis for the identification and functional characterization of the molecular components that determine the high hemocyte-specificity of expression of myticins.

Although previous studies had already indicated that myticins show remarkably high levels of expression, we here provide new evidence that each individual expresses its own repertoire of sequence variants. The combination between evidence of expression and presence/absence at the genomic level demonstrated that just a few canonical myticin clusters were present in the majority of individuals. In contrast, the vast majority of the isoforms are found with low frequency in mussel populations, to the point that we could only identify several of them in a single individual.

5.5 CONCLUSIONS

In summary, a total of 120 different myticin variants have been defined and phylogenetically analyzed. All of these variants are subject to presence/absence variation, although with different frequency. As expected, the most highly conserved residues of the mature peptide sequence, i.e., the 8 cysteines involved in the formation of the disulfide array, were mostly found to be subjected to strong negative selection, along with a few other previously unreported sites whose functional importance will need to be investigated. While this indicates that deleterious alleles are removed, whenever a non-synonymous mutation occurs in these positions, a high number of other sites (accounting for about 25% of the mature peptide sequence) were found to show signatures of positive selection, which explains the high level of intraspecific sequence diversity observed. The identification of multiple sequence variants in each individual, together with the residual presence of several pseudogenes, further suggests that the molecular diversification of myticins has been made possible by multiple independent gene duplication events.

Most certainly, the data presented in this work indicate that the 120 variants collected from 16 individuals just represent the tip of the iceberg of an underlying extreme level of sequence polymorphism that could potentially reveal several hundred unique myticin variants with follow-up analyses of individuals belonging to populations sampled in other geographical locations.

5.6 REFERENCES

1. Mitta, G.; Hubert, F.; Noël, T.; Roch, P. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur. J. Biochem. 265, 71–78 (1999). doi:10.1046/j.1432-1327.1999.00654.x.

2. Pallavicini, A.; del Mar Costa, M.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev. Comp. Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

3. Costa, MM.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev. Comp. Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

4. Vera, M.; Martínez, P.; Poisa-Beiro, L.; Figueras, A.; Novoa, B. Genomic organization, molecular diversification, and evolution of antimicrobial peptide myticin-C genes in the mussel (*Mytilus galloprovincialis*). PLoS ONE. 6, e24041 (2011). doi:10.1371/journal.pone.0024041.

5. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: A chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS ONE. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

6. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral Activity of Myticin C Peptide from Mussel: An Ancient Defense against Herpesviruses. J. Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

7. Moreira, R.; Pereiro, P.; Canchaya, C.; Posada, D.; Figueras, A.; Novoa, B. RNA-Seq in *Mytilus galloprovincialis*: Comparative transcriptomics and expression profiles among different tissues. BMC Genom. 16, 728 (2015). doi:10.1186/s12864-015-1817-5.

8. Hubert, F.; Noel, T.; Roch, P. A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur. J. Biochem. 240, 302–306 (1996). doi:10.1111/j.1432-1033.1996.0302h.x.

9. Mitta, G.; Vandenbulcke, F.; Hubert, F.; Salzet, M.; Roch, P. Involvement of mytilins in mussel antimicrobial defense. J. Biol. Chem. 275, 12954–12962 (2000). doi:10.1074/jbc.275.17.12954.

10. Sonthi, M.; Toubiana, M.; Pallavicini, A.; Venier, P.; Roch, P. Diversity of coding sequences and gene structures of the antifungal peptide mytimycin (MytM) from the Mediterranean mussel, *Mytilus galloprovincialis*. Mar. Biotechnol. 13, 857–867 (2011). doi:10.1007/s10126-010-9345-4.

11. Rosani, U.; Varotto, L.; Rossi, A.; Roch, P.; Novoa, B.; Figueras, A.; Pallavicini, A.; Venier, P. Massively parallel amplicon sequencing reveals isotype-specific variability of antimicrobial peptide transcripts in *Mytilus galloprovincialis*. PLoS ONE. 6, e26680 (2011). doi:10.1371/journal.pone.0026680.

12. Mitta, G.; Vandenbulcke, F.; Noel, T.; Romestand, B.; Beauvillain, J.C.; Salzet, M.; Roch, P. Differential distribution and defence involvement of antimicrobial peptides in mussel. J. Cell Sci. 113, 2759–2769 (2000).

13. Costa, M.M.; Prado-Alvarez, M.; Gestal, C.; Li, H.; Roch, P.; Novoa, B.; Figueras, A. Functional and molecular immune response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes against pathogen-associated molecular patterns and bacteria. Fish Shellfish Immunol. 26, 515–523 (2009). doi:10.1016/j.fsi.2009.02.001.

14. Gerdol, M.; Moreira, R.; Cruz, F.; Gómez-Garrido, J.; Vlasova, A.; Rosani, U.; Venier, P.; Naranjo-Ortiz, M.A.; Murgarella, M.; Greco, S.; Balseiro, P.; Corvelo, A.; Frias, L.; Gut, M.; Gabaldón, T.; Pallavicini, A.; Canchaya, C.; Novoa, B.; Alioto, T.S.; Posada, D.; Figueras, A. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome biology. 21, 275 (2020). doi:10.1186/s13059-020-02180-3.

15. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549 (2018). doi:10.1093/molbev/msy096.

16. Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT Suite: A web server for clustering and comparing biological sequences. Bioinformatics. 26, 680–682 (2010). doi:10.1093/bioinformatics/btq003.

17. Guindon, S.; Gascuel, O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52, 696–704 (2003). doi:10.1080/10635150390235520.

18. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. Nat. Methods. 9, 772 (2012). doi:10.1038/nmeth.2109.

19. Fu, Y.X. Linear invariants under Jukes' and Cantor's one-parameter model. J. Theor. Biol. 173, 339–352 (1995). doi:10.1006/jtbi.1995.0067.

20. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17, 754–755 (2001). doi:10.1093/bioinformatics/17.8.754.

21. Rambaut, A. FigTree, A Graphical Viewer of Phylogenetic Trees; University of Edinburgh: Edinburgh, UK. (2009).

22. Bendtsen, J.D.; Nielsen, H.; Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795 (2004). doi:10.1016/j.jmb.2004.05.028.

23. Kozlowski, L.P. IPC—Isoelectric Point Calculator. Biol. Direct. 11, 55 (2016). doi:10.1186/s13062-016-0159-9.

24. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond, S.L. Detecting Individual Sites Subject to Episodic Diversifying Selection. PLoS Genet. 8, e1002764 (2012). doi:10.1371/journal.pgen.1002764.

25. Kosakovsky Pond, S.L.; Frost, S.D. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. Mol. Biol. Evol. 22, 1208–1222 (2005). doi:10.1093/molbev/msi105.

26. Murrell, B.; Moola, S.; Mabona, A.; Weighill, T.; Sheward, D.; Kosakovsky Pond, S.L.; Scheffler, K. FUBAR: A fast, unconstrained bayesian approximation for inferring selection. Mol. Biol. Evol. 30, 1196–1205 (2013). doi:10.1093/molbev/mst030.

27. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. Mol. Biol. Evol. 35, 773–777 (2018). doi:10.1093/molbev/msx335.

28. Domeneghetti, S.; Franzoi, M.; Damiano, N.; Norante, R.; El Halfawy, N.M.; Mammi, S.; Marin, O.; Bellanda, M.; Venier, P. Structural and Antimicrobial Features of Peptides Related to Myticin C, a Special Defense Molecule from the Mediterranean Mussel *Mytilus galloprovincialis*. J. Agric. Food Chem. 63, 9251–9259 (2015). doi:10.1021/acs.jafc.5b03491.

29. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004). doi:10.1002/jcc.20084.

30. Reese, M.G.; Eeckman, F.H.; Kulp, D.; Haussler, D. Improved Splice Site Detection in Genie. J. Comp. Biol. 4, 311–323 (1997). doi:10.1089/cmb.1997.4.311.

31. Zia, A.; Moses, A.M. Towards a theoretical understanding of false positives in DNA motif finding. BMC Bioinform. 13, 151 (2012). doi:10.1186/1471-2105-13-151.

32. Bailey, T.L.; Bodén, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208 (2009). doi:10.1093/nar/gkp335.

33. Greco, S.; Gerdol, M.; Edomi, P.; Pallavicini, A. Molecular Diversity of Mytilin-Like Defense Peptides in Mytilidae (Mollusca, Bivalvia). Antibiotics. 9, 37 (2020). doi:10.3390/antibiotics9010037.

34. McInerney, J.O.; McNally, A.; O'Connell, M.J. Why prokaryotes have pangenomes. Nat. Microbiol. 2, 17040 (2017). doi:10.1038/nmicrobiol.2017.40.

35. Hirsch, C.N.; Foerster, J.M.; Johnson, J.M.; Sekhon, R.S.; Muttoni, G.; Vaillancourt, B.; Peñagaricano, F.; Lindquist, E.; Pedraza, M.A.; Barry, K.; de Leon, N.; Kaeppler, S.M.; Buell, C.R. Insights into the maize pan-genome and pan-transcriptome. Plant Cell. 26, 121–135 (2014). doi:10.1105/tpc.113.119982.

36. Conrad, D.F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T.D.; Barnes, C.; Campbell, P.; Fitzgerald, T.; Hu, M.; Ihm, C.H.; Kristiansson, K.; Macarthur, D.G.; Macdonald, J.R.; Onyiah, I.; Pang, A.W.; Robson, S.; Stirrups, K.; Valsesia, A.; Walter, K.; Wei, J.; Wellcome Trust Case Control Consortium; Tyler-Smith, C.; Carter, N.P.; Lee, C.; Scherer, S.W.; Hurles, M.E. Origins and functional impact of copy number variation in the human genome. Nature. 464, 704–712 (2010). doi:10.1038/nature08516.

37. Ramos-Onsins, S.; Aguadé, M. Molecular evolution of the Cecropin multigene family in Drosophila. functional genes vs. pseudogenes. Genetics. 150, 157–171 (1998).

38. Hanson, M.A.; Lemaitre, B.; Unckless, R.L. Dynamic Evolution of Antimicrobial Peptides Underscores Trade-Offs Between Immunity and Ecological Fitness. Front. Immunol. 10, 2620 (2019). doi:10.3389/fimmu.2019.02620.

39. Yang, Y.S.; Mitta, G.; Chavanieu, A.; Calas, B.; Sanchez, J.F.; Roch, P.; Aumelas, A. Solution structure and activity of the synthetic four-disulfide bond Mediterranean mussel defensin (MGD-1). Biochemistry. 39, 14436–14447 (2000). doi:10.1021/bi0011835.

40. Roch, P.; Yang, Y.; Toubiana, M.; Aumelas, A. NMR structure of mussel mytilin, and antiviral-antibacterial activities of derived synthetic peptides. Dev. Comp. Immunol. 32, 227–238 (2008). doi:10.1016/j.dci.2007.05.006.

41. Tincu, J.A.; Taylor, S.W. Antimicrobial Peptides from Marine Invertebrates. Antimicrob. Agents Chemother. 48, 3645–3654 (2004). doi:10.1128/AAC.48.10.3645-3654.2004.

42. Bechinger, B.; Gorr, S.U. Antimicrobial Peptides: Mechanisms of Action and Resistance. J. Dent. Res. 96, 254–260 (2017). doi:10.1177/0022034516679973.

43. Padhi, A.; Verghese, B. Molecular diversity and evolution of myticin-C antimicrobial peptide variants in the Mediterranean mussel, *Mytilus galloprovincialis*. Peptides. 29, 1094–1101 (2008). doi:10.1016/j.peptides.2008.03.007.

44. Vallecillo-Viejo, I.C.; Liscovitch-Brauer, N.; Diaz Quiroz, J.F.; Montiel-Gonzalez, M.F.; Nemes, S.E.; Rangan, K.J.; Levinson, S.R.; Eisenberg, E.; Rosenthal, J.J.C. Spatially regulated editing of genetic information within a neuron. Nucleic Acids Res. 48, 3999–4012 (2020). doi:10.1093/nar/gkaa172.

45. Moreira, R.; Pereiro, P.; Balseiro, P.; Milan, M.; Pauletto, M.; Bargelloni, L.; Novoa, B.; Figueras, A. Revealing *Mytilus galloprovincialis* transcriptomic profiles during ontogeny. Dev. Comp. Immunol. 84, 292–306 (2018). doi:10.1016/j.dci.2018.01.016.

46. de Boer, C.G.; Vaishnav, E.D.; Sadeh, R.; Abeyta, E.L.; Friedman, N.; Regev, A. Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. Nat. Biotechnol. 38, 56–65 (2020). doi:10.1038/s41587-019-0315-8.

47. Andersson, R.; Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. Nat. Rev. Genet. 21, 71–87 (2020). doi:10.1038/s41576-019-0173-8.

48. Fuda, N.J.; Ardehali, M.B.; Lis, J.T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. Nature. 461, 186–192 (2009). doi:10.1038/nature08449.

49. Savinkova, L.K.; Ponomarenko, M.P.; Ponomarenko, P.M.; Drachkova, I.A.; Lysova, M.V.; Arshinova, T.V.; Kolchanov, N.A. TATA box polymorphisms in human gene promoters and associated hereditary pathologies. Biochemistry, 74, 117–129 (2009). doi:10.1134/s0006297909020011.

50. Hasegawa, Y.; Struhl, K. Promoter-specific dynamics of TATA-binding protein association with the human genome. Genome Res. 29, 1939–1950 (2019). doi:10.1101/gr.254466.119.

51. Saxonov, S.; Berg, P.; Brutlag, D.L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. USA. 103, 1412–1417 (2006). doi:10.1073/pnas.0510310103.

52. Juven-Gershon, T.; Kadonaga, J.T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev. Biol. 339, 225–229 (2010). doi:10.1016/j.ydbio.2009.08.009.

53. Cramer, P. Organization and regulation of gene transcription. Nature. 573, 45–54 (2019). doi:10.1038/s41586-019-1517-4.

54. Müller, F.; Tora, L. Chromatin and DNA sequences in defining promoters for transcription initiation. Biochim. Biophys. Acta. 1839, 118–128 (2014). doi:10.1016/j.bbagrm.2013.11.003.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSIONS

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 6: General discussion and conclusions

6.1 General discussion

Mediterranean mussel research has gained importance in the last two decades, both for its environmental importance and immune features. The recent sequencing of its genome [1] revealed a complex pan-genomic architecture, being 25% of their genes, entirely dispensable, and therefore subject to presence/absence variation (PAV). Because of that, each individual shows their own repertoire of genes, and therefore different expression profiles. In the first work of this thesis **("Chapter 2: High individual variability in the transcriptomic response of Mediterranean mussels to *Vibrio* reveals the involvement of myticins in tissue injury"**), we could verify major differences in transcriptomic response of different mussels to the same stimulus, both in terms of the number of expressed genes as well as expression magnitude. This fact suggests that a careful experimental design is necessary, especially with non-model species, as their individual genomes, and therefore transcriptomes can be quite diverse. Additionally, some common experimental procedures in bivalves such as shell notching or the administration of substances through an injection into the adductor muscle could trigger unexpected reactions such as unspecific inflammatory responses. In this way, the study of the response to a bacterial infection (by *Vibrio splendidus*) showed an unpredicted up-regulation of certain AMPs in control individuals. This is the case of some myticins, that have been classically defined as antimicrobial [2,3] but other functions cannot be ignored since, for example, the chemotactic capacity of myticin C has been defined [4].

Because of that, myticin C could act as a chemokine preparing mussels for future pathogenic processes after a danger signal, instead of acting by killing the pathogens in a strict way.

After knowing the transcriptomical response of individual mussels before and after a bacterial infection, we were interested in studying the immune responses of mussel hemocytes after two consecutive encounters with the same pathogen (*V. splendidus*). This work constitutes the next chapter of this thesis **("Chapter 3: Immune tolerance in *Mytilus galloprovincialis* hemocytes after repeated contact with *Vibrio splendidus*"**) and it allowed to increase the knowledge in mussel hemocytes of the new emerging concept of "trained immunity" [5,6,7], according to which innate immune cells can become reprogrammed to develop immunological memory after previous encounters with non-self-molecules. One of the most studied bivalves in these terms is the oyster *Crassostrea gigas*. It has been demonstrated how a priming with poly I:C provides protection from subsequent exposure to OsHV-1 [8], and this protection can be extended up to 5 months [9,10]. Moreover, offspring produced from poly I:C-treated parents doubled their chances of surviving exposures to OsHV-1 [11]. The analysis of the response of mussel hemocytes after two successive exposures to *V. splendidus* suggests that, after the second contact with the bacteria, immune cells attempted to control and resolve the inflammatory response to avoid subsequent DNA damage and cell death. There appeared a tightly regulated response shifting from a pro-inflammatory response to an anti-inflammatory and probably regenerative phenotype. It could be interesting to know if this modified response is pathogen-specific or not, or how long the modifications generated by the priming process are maintained. In any case, the mechanism of trained immunity continues to be a challenge, especially in non-model species such as mussels, that could be very valuable for comparative immunology studies.

At this point, it seemed interesting to study how the mussel immune cells could react after an *in vitro* exposure to one of their most expressed peptides, the myticin C. The **Chapter 4** of this thesis **("Transcriptomic analysis reveals the wound healing activity of mussel myticin C"**) reveals that myticin C has such an effect on hemocytes that causes

changes in their expression profile and mobility behavior. In terms of expression, a great number of genes directly related to the actin cytoskeleton were up-modulated by the peptide, which was supported by functional experiments, revealing an enhanced ability of the cells to migrate after a myticin C-treatment. This is also in line with what Balseiro et al., defined in 2011 [4], as well as the results obtained in the chapter 2 of this thesis, which pointed to the chemotactic activity of the molecule. This activity is key in biological processes such as cell migration, which in turn determines immune and regeneration pathways [12,13,14]. Cell migration is the first action that happens after an injury. The production of growth factors and chemokines attract new cells in charge of removing debris and stimulate angiogenesis, as well as extracellular matrix production [15]. In mussels, hemocytes would be attracted by chemokine-like molecules, such as myticin C, and would play their role in the damage resolution. At the same time, the increase of hemocytes would enhance the production and release of more myticin C in the damage area. But one question remains unsolved regarding the effects of tissue damage on the animal. We have not yet been able to know if cell proliferation and hematopoiesis occur after the damage. Future work to elucidate the existence or not of these mechanisms would be very interesting and would represent a great advance in the knowledge of the immunity of this species.

A special mention deserves another finding of this chapter 4, based precisely on the proposed theory that myticin C is a cytokine-like protein. Using cellular and animal models such as HaCaT cells and wild-type zebrafish larvae, it could be possible to confirm the capability of myticin C to accelerate processes of regeneration. It was observed a significant effect in the closing of the cell monolayer (after producing a disruption) and the renewal of the tissue (after making a cut in the larvae tail), so we are dealing with a peptide that could have biotechnological applications.

These studies have made possible to observe the wide functional possibilities of myticins. This fact, together with preliminary studies indicating a high variability of myticin [16,17,18] and the great mussel genome complexity recently revealed [1], made interesting to make a study of the variability of the gene (**“Chapter 5: Comparative**

genomics reveals a significant sequence variability of myticin genes in *Mytilus galloprovincialis*"). In this work, a total of 120 different myticin variants have been found in 16 different mussel genomes [1]. All of these variants are subject to presence/absence variation, so myticins belong to the group of *dispensable* genes of mussel genome. As expected, the most highly conserved residues of the mature peptide sequence, i.e., the 8 cysteines involved in the formation of the disulfide array, were mostly found to be subjected to strong negative selection. Additionally, a high number of other sites (accounting for about 25% of the mature peptide sequence) were found to show signatures of positive selection, which explains the high level of intraspecific sequence diversity observed. The identification of multiple sequence variants in each individual, together with the residual presence of several pseudogenes, further suggests that the molecular diversification of myticins has been made possibly by multiple independent gene duplication events. The data presented in this last work indicate an extreme level of sequence polymorphism that could potentially reveal several hundred unique myticin variants. Studies that include individuals from other parts of the world may serve to reveal further variability. This study lays the basis for the future development of a large database of this or other peptides, since it has become evident both with the recent sequencing of the genome [1] and in the present thesis, the very high levels of variability of some genes in the mussel *M. galloprovincialis*.

6.2 REFERENCES

1. Gerdol, M.; Moreira, R.; Cruz, F.; Gómez-Garrido, J.; Vlasova, A.; Rosani, U.; Venier, P.; Naranjo-Ortiz, M.A.; Murgarella, M.; Greco, S.; Balseiro, P.; Corvelo, A.; Frias, L.; Gut, M.; Gabaldón, T.; Pallavicini, A.; Canchaya, C.; Novoa, B.; Alioto, T.S.; Posada, D.; Figueras, A. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome Biol. 21, 275 (2020). doi:10.1186/s13059-020-02180-3.

2. Martinez-Lopez, A.; Encinar, J.A.; Medina-Gali, R.M.; Balseiro, P.; Garcia-Valtanen, P.; Figueras, A.; Novoa, B.; Estepa, A. pH-dependent solution structure and activity of a reduced form of the host-defense peptide myticin C (Myt C) from the mussel *Mytilus galloprovincialis*. Mar Drugs. 11, 2328–2346 (2013). doi:10.3390/md11072328.

3. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral activity of myticin C peptide from mussel: an ancient defense against herpesviruses. J Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

4. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS One. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

5. Netea, M.G.; Quintin, J.; Van der Meer, J.W.M. Trained immunity: a memory for innate host defense. Cell Host Microbe. 9, 355–61 (2011). doi:10.1016/j.chom.2011.04.006.

6. Quintin, J.; Cheng, S.-C.; Van der Meer, J.W.M.; Netea, M.G. Innate immune memory: towards a better understanding of host defense mechanisms. Curr Opin Immunol. 29, 1–7 (2014). doi:10.1016/j.coi.2014.02.006.

7. Gourbal, B.; Pinaud, S.; Beckers, G.J.M.; Van der Meer, J.W.M.; Conrath, U.; Netea, M.G. Innate immune memory: an evolutionary

perspective. Immunol Rev. 283, 21–40 (2018). doi:10.1111/imr.12647.

8. Green, T.J.; Montagnani, C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus (OsHV-1 μvar). Fish Shellfish Immunol. 35, 382–8 (2013). doi:10.1016/j.fsi.2013.04.051.

9. Lafont, M.; Petton, B.; Vergnes, A.; Pauletto, M.; Segarra, A.; Gourbal, B.; Montagnani, C. Long-lasting antiviral innate immune priming in the Lophotrochozoan Pacific oyster, *Crassostrea gigas*. Sci Rep. 7, 13143 (2017). doi:10.1038/s41598-017-13564-0.

10. Lafont, M.; Vergnes, A.; Vidal-Dupiol, J.; de Lorgeril, J.; Gueguen, Y.; Haffner, P.; Petton, B.; Chaparro, C.; Barrachina, C.; Destoumieux-Garzon, D.; Mitta, G.; Gourbal, B.; Montagnani, C. A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, *Crassostrea gigas*. mBio. 10, e02777-19 (2020). doi: 10.1128/mBio.02777-19.

11. Green, T.J.; Helbig, K.; Speck, P.; Raftos, D.A. Primed for success: oyster parents treated with poly(I:C) produce offspring with enhanced protection against Ostreid herpesvirus type I infection. Mol Immunol. 78, 113–20 (2016). doi:10.1016/j.molimm.2016.09.002.

12. Collins, C.; Nelson, W.J. Running with neighbors: Coordinating cell migration and cell-cell adhesion. Curr Opin Cell Biol. 36, 62–70 (2015). doi:10.1016/j.ceb.2015.07.004.

13. Huse, M. Mechanical forces in the immune system. Nat Rev Immunol. 17, 679–690 (2017). doi:10.1038/nri.2017.74.

14. Bellavia, G.; Fasanaro, P.; Melchionna, R.; Capogrossi, M.C.; Napolitano, M. Transcriptional control of skin reepithelialization. J Dermatol Sci. 73, 3–9 (2014). doi:10.1016/j.jdermsci.2013.08.007.

15. Kasuya, A.; Tokura, Y. Attempts to accelerate wound healing. J Dermatol Sci. 76, 169–172 (2014). doi:10.1016/j.jdermsci.2014.11.001.

16. Pallavicini, A.; Costa M.M.; Gestal, C.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

17. Costa, M.M.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

18. Vera, M.; Martínez, P.; Poisa-Beiro, L.; Figueras, A.; Novoa, B. Genomic organization, molecular diversification, and evolution of antimicrobial peptide myticin-C genes in the mussel (*Mytilus galloprovincialis*). PLoS ONE. 6, e24041 (2011). doi:10.1371/journal.pone.0024041.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

6.3 CONCLUSIONS

1. Mussel individual transcriptomes are quite diverse, therefore, the design of the experiments must take into account this individual variability, especially with non-model species.
2. Mussels respond to danger and injury signals by expressing antimicrobial peptides such as myticin C, which modulates the response of hemocytes to subsequent infections.
3. A reprogramming in mussel gene expression has been demonstrated in a situation of reinfection with *Vibrio splendidus*. This modification of the immune response resulted in the control of inflammation, as well as an immune tolerance phenotype.
4. Myticin C causes changes in the gene expression profile of hemocytes, changing their state of activation and mobility behavior.
5. Myticin C could have a role in tissue damage accelerating the processes of regeneration, both in an *in vitro* human skin model (HaCaT) and an *in vivo* regeneration model such as zebrafish.
6. Myticins show a great variability, having identified a total of 120 different variants that were subject to presence/absence variation.
7. The availability of genome and transcriptome of the same individual allowed to confirm that no RNA editing processes are the responsible of such a variability, and the myticin diversity have a genomic origin.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

RESUMO E CONCLUSIÓNS

RESUMO

Introdución

O mexillón mediterráneo (*Mytilus galloprovincialis*) é un molusco bivalvo que vive fixo nas rocas e cuxo modo de vida é a filtración. A súa distribución e moi extensa e a produción global alcanza os 2 millóns de toneladas ao ano [1]. Aínda que os valores produtivos de cada país son aproximados debido á ambigüidade na identificación das diferentes especies de mexillón, o principal país produtor é China (arredor de un millón de toneladas ao ano), mentres Chile e España producen unhas 230.000 toneladas. Estes valores tan altos de produción están relacionados cun incremento continuo da poboación mundial dende 1950, que está previsto que alcance os 10 mil millóns de habitantes no ano 2050 [2].

Además dos datos produtivos, moi importantes no que a dispoñibilidade de alimento e desenvolvemento económico se refire, o mexillón mediterráneo ten un gran valor biolóxico. Especialmente porque está presente en todo o planeta, sendo considerada unha especie invasora [3]. A súa extensa distribución, xunto co feito xa citado de que se trata dunha especie inmóbil e filtrante, fai que estes animais sexan interesantes como marcadores ecolóxicos de contaminación [4,5].

Outro aspecto importante polo que gañou importancia o estudo do mexillón mediterráneo é a súa particular resistencia ás enfermidades. A incapacidade para moverse e a entrada a través das branquias de todo tipo de partículas, podería levar a pensar que estes animais están en

condicións de vulnerabilidade debido á súa exposición continua a axentes patóxenos e contaminantes. Non obstante, non se rexistraron mortalidades masivas na natureza [6,7], algo que ocorreu con outras especies da clase Bivalvia como as ostras [8] ou as ameixas [9] coas que comparte hábitat. A explicación desta gran capacidade para resistir condicións adversas no ambiente pode esconderse nas incógnitas que aínda quedan por estudar do complexo sistema inmune destes animais. A pesar de carecer de inmunidade adaptativa, como ocorre entre os invertebrados, teñen unha poderosa inmunidade innata e combaten as infeccións a través dos compoñentes e mecanismos celulares que a integran.

Os hemocitos son as células inmunes centrais dos mexillóns. Estas células recoñecen patróns moleculares conservados asociados a patóxenos (PAMPs) a través de receptores de recoñecemento de patróns (PRRs), así como patróns moleculares asociados a perigo / dano (DAMPs). Despois dos procesos de recoñecemento, estas células actúan mediante quimiotaxe (recrutamento de células inmunes ao lugar danado), encapsulación (internalización de partículas esóxenas), actividade fagocítica (internalización para a posterior degradación de partículas esóxenas) e liberación de especies reactivas de osíxeno e radicais de nitróxeno (claves no proceso de destrución de patóxenos), así como a activación de vías de sinalización intracelulares para desencadear finalmente a síntese de efectores antimicrobianos.

Unha característica moi importante do xenoma destes animais é a súa gran variabilidade. Sábese que o 25% dos xenes son "prescindibles" e, polo tanto, poden estar presentes ou non nun individuo específico (un fenómeno coñecido como variación de presenza / ausencia ou PAV) [10]. Así, xenes de recoñecemento como os que conteñen o dominio C1q, lectinas, FREPs (*fibrinogen-related proteins*) ou TLRs (*Toll-like receptors*) constitúen familias de proteínas moi variables e especialmente expandidas nos mexillóns [11,12,13]. Do mesmo xeito, as moléculas efectoras como os péptidos antimicrobianos (AMPs) tamén se caracterizaron por mostrar unha gran diversidade nestes animais [14,15,16,17,18,19].

Os AMPs son moléculas que se caracterizan por ter unha estrutura de aminoácidos rica en cisteínas e con propiedades antimicrobianas. Algúns destes péptidos descritos nos mexillóns son as defensinas [14], as mitilinas [15], as mitimicinas [16] e as miticinas [17]. Estas moléculas están formadas por un péptido sinal, seguido dun péptido maduro (que contén residuos de cisteínas) e un extremo C-terminal.

De entre os AMPs mencionados, as miticinas destacan pola súa enorme variabilidade e diversidade funcional. En traballos anteriores identificáronse as isoformas A, B e C [17,18], pero as técnicas de secuenciación masiva que comezaron a implementarse nos últimos anos permitiron descubrir unha diversidade interindividual moito maior, outorgándolles un gran interese dende un punto de vista xenómico [18,19]. Ademais, no que respecta á súa función, non só mostran as características antimicrobianas que os definen [20], senón que se descubriron outras funcións. Por exemplo, sábese que a miticina C ten una importante capacidade quimiotáctica, razón pola que se considerou como a primeira molécula de tipo citoquina identificada nos bivalvos [21].

Ademais da variabilidade xenómica e a gran diversidade molecular destes animais para afrontar a invasión de patóxenos, outro factor que pode ser clave na resposta inmune do mexillón é a memoria da resposta inmune innata. Este concepto coñécese como "*priming immunity*" ou "*trained immunity*" e contempla a posibilidade de que os invertebrados, carentes de anticorpos, vexan aumentada a súa resposta inmune en situacións de contacto repetido cos mesmos ou diferentes axentes patóxenos [22,23,24]. Este mecanismo axudaría a comprender como os invertebrados poden facer fronte ás infeccións dun xeito tan exitoso. Aínda que hai que ter en conta que esta memoria é moito menos específica e produce menos amplificación da resposta que a inmunidade adaptativa dos animais vertebrados.

Entre os invertebrados identificáronse moitos exemplos de "*trained immunity*" [23,24,25]. O bivalvo máis estudado ata a data ao respecto é a ostra (*Crassostrea gigas*), na que se viu como o poly I:C (*polyinosinic:polycytidylic acid*), que é un inmunoestimulante que simula infeccións virais, desencadea unha resposta antiviral e

protección contra infeccións posteriores co herpesvirus da ostra (OsHV-1) [26]. Confirmouse que esta protección pode durar ata 5 meses [27,28] e, ademais, a descendencia de proxenitores estimulados con poly I:C mostra o dobre de opcións de sobrevivir despois da exposición ao virus [29]. Non hai moitos estudos deste mecanismo en mexillóns, pero a idea de que a inmunidade innata destes animais se vería modificada despois de repetidas estimulacións cun patóxeno foi obxecto de estudo nesta tese.

O desenvolvemento de técnicas de secuenciación masiva e a súa implementación en animais non modelo como o mexillón xeraron moita información. En particular, a mencionada secuenciación do xenoma [10], así como diferentes estudos transcriptómicos permitiron enriquecer as bases de datos con secuencias de interese. Segundo isto, hai información dispoñible en termos de expresión de varios tecidos (manto, músculo, branquias, hemocitos...) [30,31], etapas de desenvolvemento [32] ou en situacións de infección [33]. Esta posibilidade de coñecer o perfil transcriptómico completo dun organismo ofrécena técnicas como o *RNA-Seq*, que é a tecnoloxía máis empregada neste campo. Esta técnica é fundamental na presente tese e permitiunos coñecer a expresión de numerosos xenes, ampliando o coñecemento da resposta inmune do mexillón.

Obxectivos

O obxectivo principal desta tese é afondar no coñecemento da resposta inmune do mexillón. En particular, aumentar a información xenética e transcriptómica, así como continuar investigando a variabilidade funcional de xenes como os AMPs.

En particular:

1. Aumentar a información sobre como os hemocitos responden a unha infección bacteriana (con *Vibrio splendidus*), centrando o estudo na resposta individual.

2. Estudar a resposta transcriptómica dos hemocitos a dúas exposicións co mesmo patóxeno, buscando responder preguntas sobre a existencia dunha memoria innata nos bivalvos.
3. Analizar, dende un punto de vista transcriptómico, o efecto directo da miticina C sobre as células inmunes do mexillón, procurando ampliar o coñecemento sobre as posibilidades funcionais do péptido.
4. Estudar a secuencia das miticinas para determinar a orixe da súa gran variabilidade, así como a información evolutiva desta familia de péptidos antimicrobianos, que son fundamentais para comprender a gran resistencia destes animais no medio natural.

Discusión

A recente secuenciación do xenoma do mexillón [10] revelou a gran complexidade molecular destes animais. A arquitectura deste panxenoma ten ao redor dun 25% de xenes "prescindibles" suxeitos a variación de presenza/ausencia (PAV). Por esta razón, cada individuo mostra o seu propio repertorio de xenes e, polo tanto, diferentes perfís de expresión. No primeiro traballo desta tese **("Capítulo 2: A gran variabilidade individual na resposta transcriptómica do mexillón mediterráneo a *Vibrio* revela a participación das miticinas nas lesións dos tecidos"**), puidemos comprobar diferenzas importantes na resposta transcriptómica de diferentes mexillóns ante o mesmo estímulo. Este feito suxire que é necesario un deseño experimental coidadoso, especialmente en especies non modelo, xa que os seus xenomas e, polo tanto, os transcriptomas, poden ser bastante diversos. Neste sentido, a natureza limitante dalgúns materiais biolóxicos, como os hemocitos, levou ao uso de grupos de individuos na maioría dos estudos transcriptómicos realizados en bivalvos [34,35]. Esta solución axuda a eliminar as diferenzas individuais e permite centrar a análise na condición ou tratamento en estudo. Pero, no caso de querer coñecer respostas individuais específicas, como é o caso do capítulo 2 desta tese, a estratexia debe ter en conta a mostraxe individual dos animais.

Neste capítulo puidemos observar como a resposta de cada mexillón ao mesmo estímulo era bastante diverxente. Cada mexillón mostrou o seu propio repertorio de xenes regulados tanto nos controis como nos infectados con *Vibrio splendidus*, mostrando a gran variabilidade xenética xa mencionada anteriormente.

Por outra banda, no que se refire aos estímulos probados, o dano nos tecidos (representado pola inxección de auga de mar nos controis) provocou unha resposta relacionada coa proliferación e migración celular. Esta resposta transcriptómica viuse apoiada por estudos funcionais de citometría de fluxo nos que foi posible observar como o dano nos tecidos pode preparar os hemocitos para una situación posterior de infección. Isto é importante xa que algúns procedementos experimentais comúns en bivalvos, como perforar a cuncha ou administrar substancias por inxección no músculo adutor, poderían desencadear reaccións inesperadas e respostas inflamatorias inespecíficas.

Ademais, no que a resposta a una infección se refire, puidose ver una importante sobreexpresión de xenes antimicrobianos e reguladores de procesos inflamatorios como o *IRG1* (*Immunoresponsive gene 1*), *SOCS2* (*Suppressor Of Cytokine Signaling 2*) e *Myd88* (*Myeloid Differentiation Primary Response 88*). Non obstante, entre os xenes máis expresados despois da infección, non se atoparon AMPs, algúns dos cales sí mostraron un incremento inesperado da expresión nos individuos control. Este é o caso de algunas miticinas, clasicamente definidas como antimicrobianas [20], pero que poden mostrar outras funcións. Por exemplo, a miticina C, cuxa capacidade quimiotáctica abriu a porta a moitas outras propiedades biolóxicas [21]. Por este motivo, na presente tese considérase a hipótese de que a miticina C podería actuar como unha quimioquina que prepara o mexillón para situacións inmunes esixentes despois dun sinal de perigo, no canto de actuar estritamente na destrución de axentes patóxenos.

Despois de coñecer a resposta transcriptómica dos mexillóns individuais antes e despois da infección bacteriana, estudouse a resposta dos hemocitos do mexillón despois de dous encontros consecutivos co mesmo patóxeno (*V. splendidus*). Este traballo constitúe o seguinte

capítulo desta tese **("Capítulo 3: Tolerancia inmune dos hemocitos de *Mytilus galloprovincialis* despois dun contacto repetido con *Vibrio splendidus*"**) e permitiu aumentar o coñecemento nas células inmunes do mexillón do concepto emerxente "*trained immunity*". [22]. Este concepto baséase na idea de que as células inmunes innatas poden reprogramarse para desenvolver memoria inmune despois de encontros previos con moléculas non propias. Como xa se mencionou na introdución, un dos bivalvos máis estudados nestes termos é a ostra *Crassostrea gigas*, na que se demostrou que o tratamento co inmunoestimulante poly I:C proporciona protección contra OsHV-1 [26,27, 28.29].

No caso dos mexillóns, a magnitude da resposta despois da primeira infección é moito maior que a segunda. En concreto, a primeira estimulación desencadeou unha modulación positiva da resposta inflamatoria e da migración celular. Non obstante, despois do segundo contacto coa bacteria, a resposta indica o control e resolución da resposta inflamatoria para evitar danos posteriores no ADN e morte celular. Observouse unha regulación estrita que pasou dunha situación de inflamación a un fenotipo antiinflamatorio e probablemente rexenerativo. Este cambio de resposta é importante, xa que un estado inflamatorio permanente sería letal para estes animais que están permanentemente expostos á entrada de patóxenos no seu interior.

Ademais da inflamación, outros procesos inmunes importantes como a produción de especies reactivas de osíxeno (ROS) ou a apoptose controláronse despois da reinfección. En ambos procesos observouse unha inhibición, polo que se pode concluír que dúas exposicións sucesivas a doses subletais de *Vibrio splendidus* orixinan unha resposta dirixida a tolerar a segunda infección e minimizar os danos nos tecidos.

Aínda quedan moitas incógnitas por resolver neste campo da memoria inmune innata, como por exemplo, se esta resposta é específica de patóxeno ou canto tempo se conservan as modificacións xeneradas. En calqueira caso, coñecer estos mecanismos en especies tan resistentes como o mexillón pode supoñer un gran valor no desenvolvemento de estudios de inmunoloxía comparada.

Neste punto decidimos estudar como as células inmunes do mexillón poderían reaccionar despois da exposición *in vitro* a un dos péptidos máis expresados do transcriptoma dos hemocitos do mexillón, a miticina C. O **Capítulo 4** desta tese ("**Análises transcriptómicas revelan a actividade rexeneradora da miticina C do mexillón**") revela que a miticina C provoca cambios no perfil de expresión e na mobilidade dos hemocitos. En termos de expresión, un gran número de xenes relacionados co citoesqueleto de actina moduláronse e sobreexpresáronse debido ao tratamento co péptido. Diversos experimentos funcionais apoiaron estos resultados e revelaron que as células migran máis despois do tratamento con miticina C, coincidindo con investigacións anteriores [21], así como cos resultados obtidos no capítulo 2 desta tese, que apuntaron á actividade quimiotáctica da molécula. Esta actividade é clave en procesos biolóxicos como a migración celular, que á súa vez condiciona as rutas de inmunidade e rexeneración [36,37,38]. A migración celular é a primeira acción que se produce despois dunha lesión. A produción de factores de crecemento e quimioquinas atrae novas células encargadas de eliminar os restos celulares e tisulares e estimular a anxioxénese, así como a produción de matriz extracelular [39]. Nos mexillóns, os hemocitos ven modificada a súa mobilidade por moléculas de tipo quimioquina, como a miticina C, que os atrae a áreas danadas onde desempeñarían o seu papel na resolución do problema. Ao mesmo tempo, o aumento do número de hemocitos incrementaría a produción e liberación de máis miticina C na área danada. Usando modelos celulares e animais como as células HaCaT e as larvas de peixe cebra, foi posible probar a capacidade da miticina C para acelerar os procesos de rexeneración. Observouse un efecto significativo no peche do tapiz celular (despois de producir unha rotura da capa celular) e na renovación do tecido (despois de facer un corte na cola das larvas). Segundo isto, o péptido acelera os procesos de rexeneración en células tanto de bivalvos como peixes e mamíferos. Todos estes resultados avalan a xa proposta teoría de que a miticina C é unha proteína de tipo citoquina exclusiva dos mexillóns e que podería ter potencial biotecnolóxico.

Estes estudos permitiron descubrir as amplas posibilidades funcionais das miticinas. Ademais, estudos preliminares que indican

unha alta variabilidade a nivel de secuencia do péptido [18,19] e a recente publicación do xenoma do mexillón [10] deron lugar ao último capítulo desta tese, que tivo como obxectivo o estudo da variabilidade xénica das miticinas (**"Capítulo 5: Estudos de xenómica comparada revelan unha gran variabilidade da secuencia dos xenes da miticina en *Mytilus galloprovincialis*"**). Neste traballo atopáronse un total de 120 variantes de miticina en 16 xenomas diferentes de mexillón [10]. Todas estas variantes están suxeitas a variacións de presenza/ausencia, polo que as miticinas pertencen ao grupo de xenes "prescindibles" do xenoma do mexillón. É por iso que cada individuo dos 16 analizados mostrou o seu propio repertorio de formas, das que aproximadamente a metade eran exclusivas e non se atoparon en ningún dos outros individuos. A identificación de varias variantes de secuencia en cada individuo, xunto coa presenza residual de varios pseudoxenes, suxire ademais que a diversificación molecular das miticinas foi posible grazas a múltiples eventos de duplicación de xenes independentes.

En termos funcionais, estes péptidos ricos en cisteínas (os AMPs) adoitan ter unha natureza catiónica e anfipática que facilita a súa interacción electrostática coas bacterias. Neste caso, o punto isoeléctrico e a carga neta das miticinas foron sorprendentemente baixos, un aspecto que se pode explicar pola diversidade funcional do péptido, que, como se ve nos capítulos 2 e 4 desta tese, estaría máis ligada ás funcións tipo citoquina e rexeneradora que á estrita destrución de axentes patóxenos.

Aínda así, as cisteínas son fundamentais para explicar a función destes péptidos. Isto demostrouse coa análise de selección positiva e negativa realizada, na que se viu que os residuos máis conservados da secuencia peptídica madura son as 8 cisteínas implicadas na formación de pontes disulfuro e que están principalmente sometidas a un forte selección negativa. Ademais, outros aminoácidos importantes para manter a estrutura secundaria da proteína tamén están sometidos a selección purificadora (unha serina e unha treonina da rexión da hélice alfa e unha glicina que une as dúas láminas beta que compoñen as miticinas). Pola contra, atopouse un gran número de aminoácidos (que representan aproximadamente o 25% da secuencia peptídica madura) que mostraron sinaturas de selección positiva ou diversificadora, o que

explica o alto nivel de variabilidade observado. Descoñécese a orixe de toda esta diversidade, pero a dispoñibilidade do xenoma e do transcriptoma do mesmo individuo permitiu unha comparación entre o ADN xenómico e o ARN mensaxeiro. Isto permitiunos observar que non hai procesos de edición de ARN detrás da xeración dun número tan alto de variantes, senón que a variabilidade das miticinas ten unha orixe completamente xenómica.

No que se refire á expresión, as miticinas son un dos xenes máis expresados nos mexillóns. Isto suxire unha forte regulación da transcrición, polo que os elementos reguladores do xene poden incluír promotores e potenciadores fortes e conservados que regulan dita expresión. Neste sentido, o estudo da rexión promotora permitiu identificar ata 10 motivos diferentes, moi conservados en número e disposición, apoiando resultados anteriores nos que se observaron valores de expresión moi altos.

Os datos presentados neste último traballo indican un nivel extremo de polimorfismo da secuencia, deixando claro que, no caso de analizar individuos pertencentes a poboacións doutras localizacións xeográficas, poderían descubrirse varios centos de variantes únicas de miticina.

Este traballo establece as bases para o desenvolvemento dunha gran base de datos que recolla as variantes deste e outros xenes, xa que tanto a recente secuenciación do xenoma como a presente tesis deixan patente os altos niveis de variabilidade dalgúns xenes do mexillón *Mytilus galloprovincialis*.

Referencias

1. FAO. FAO Home Page. Available online: http://www.fao.org/home/en/ (accessed on 17 August 2020).
2. United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019, custom data acquired via website (accessed on 19 August 2020).

3. Bonham, V. *Mytilus galloprovincialis*. Invasive Species Compendium (CAB, Wallingford, UK). (2017).

4. Goldberg, E.D. The mussel watch — A first step in global marine monitoring. Mar Pollut Bull. 6, 111 (1975).

5. Fiorito, F.; Amoroso, M.G.; Lambiase, S.; Serpe, F.P.; Bruno, T.; Scaramuzzo, A.; Maglio, P.; Fusco, G.; Esposito, M. A relationship between environmental pollutants and enteric viruses in mussels (*Mytilus galloprovincialis*). Environ Res. 169, 156–162 (2019). doi:10.1016/j.envres.2018.11.001.

6. Romero, A.; del Mar 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. 108, 149–163 (2014). doi:10.3354/dao02701.

7. Benabdelmouna, A.; Garcia, C.; Ledu, C.; Lamy, P.; Maurouard, E.; Dégremont, L. Mortality investigation of *Mytilus edulis* and *Mytilus galloprovincialis* in France: An experimental survey under laboratory conditions. Aquaculture. 495, 831–841 (2018). doi:10.1016/j.aquaculture.2018.06.075.

8. Segarra, A.; Pepin, J.F.; Arzul, I.; Morga, B.; Faury, N.; Renault, T. Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res. 153, 92–99 (2010). doi:10.1016/j.virusres.2010.07.011.

9. Garcia, C.; Haond, C.; Chollet, B.; Nerac, M.; Omnes, E.; Joly, J.P.; Dubreuil, C.; Serpin, D.; Langlade, A.; Le Gal, D.; Terre-Terrillon, A.; Courtois, O.; Guichard, B.; Arzul, I. Descriptions of *Mikrocytos veneroïdes* n. sp. and *Mikrocytos donaxi* n. sp. (Ascetosporea: Mikrocytida: Mikrocytiidae), detected during important mortality events of the wedge clam *Donax trunculus Linnaeus* (Veneroida: Donacidae), in France between 2008 and 2011. Parasit Vectors. 11, 119 (2018). doi:10.1186/s13071-018-2692-0.

10. Gerdol, M.; Moreira, R.; Cruz, F.; Gómez-Garrido, J.; Vlasova, A.; Rosani, U.; Venier, P.; Naranjo-Ortiz, M.A.; Murgarella, M.; Greco, S.; Balseiro, P.; Corvelo, A.; Frias, L.; Gut, M.; Gabaldón, T.; Pallavicini, A.; Canchaya, C.; Novoa, B.; Alioto, T.S.; Posada, D.; Figueras, A. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome Biol. 21, 275 (2020). doi:10.1186/s13059-020-02180-3.

11. Gerdol, M.; Manfrin, C.; De Moro, G.; Figueras, A.; Novoa, B.; Venier, P.; Pallavicini, A. 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 (2011). doi:10.1016/j.dci.2011.01.018.

12. Romero, A.; Dios, S.; Poisa-Beiro, L.; Costa, M.M.; 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. 35, 334–344 (2011). doi:10.1016/j.dci.2010.10.007.

13. Toubiana, M.; Gerdol, M.; Rosani, U.; Pallavicini, A.; Venier, P.; Roch, P. Toll-like receptors and MyD88 adaptors in Mytilus: complete cds and gene expression levels. Dev Comp Immunol. 40, 158–66 (2013). doi:10.1016/j.dci.2013.02.006.

14. Hubert, F.; Noel, T.; Roch, P. A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem. 240, 302–306 (1996). doi:10.1111/j.1432-1033.1996.0302h.x.

15. Mitta, G.; Hubert, F.; Dyrynda, E.A.; Boudry, P.; Roch, P. Mytilin B and MGD2, two antimicrobial peptides of marine mussels: gene structure and expression analysis. Dev Comp Immunol. 24, 381–393 (2000). doi:10.1016/s0145-305x(99)00084-1.

16. Sonthi, M.; Toubiana, M.; Pallavicini, A.; Venier, P.; Roch, P. Diversity of coding sequences and gene structures of the antifungal peptide mytimycin (MytM) from the Mediterranean mussel, *Mytilus*

galloprovincialis. Mar Biotechnol. 13, 857–867 (2011). doi:10.1007/s10126-010-9345-4.

17. Mitta, G.; Hubert, F.; Noël, T.; Roch, P. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur J Biochem. 265, 71–78 (1999). doi:10.1046/j.1432-1327.1999.00654.x.

18. Pallavicini, A.; Costa M.M.; Gestal, C.; Gestal, C.; Dreos, R.; Figueras, A.; Venier, P.; Novoa, B. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 32, 213–226 (2008). doi:10.1016/j.dci.2007.05.008.

19. Costa, M.M.; Dios, S.; Alonso-Gutierrez, J.; Romero, A.; Novoa, B.; Figueras, A. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol. 33, 162–170 (2009). doi:10.1016/j.dci.2008.08.005.

20. Novoa, B.; Romero, A.; Álvarez, Á.L.; Moreira, R.; Pereiro, P.; Costa, M.M.; Dios, S.; Estepa, A.; Parra, F.; Figueras, A. Antiviral Activity of Myticin C Peptide from Mussel: An Ancient Defense against Herpesviruses. J Virol. 90, 7692–7702 (2016). doi:10.1128/JVI.00591-16.

21. Balseiro, P.; Falcó, A.; Romero, A.; Dios, S.; Martínez-López, A.; Figueras, A.; Estepa, A.; Novoa, B. *Mytilus galloprovincialis* myticin C: A chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS ONE. 6, e23140 (2011). doi:10.1371/journal.pone.0023140.

22. Netea, M.G.; Quintin, J.; Van der Meer, J.W.M. Trained immunity: a memory for innate host defense. Cell Host Microbe. 9, 355–61 (2011). doi:10.1016/j.chom.2011.04.006.

23. Kurtz, J.; Franz, K. Evidence for memory in invertebrate immunity. Nature. 425, 37–38 (2003). https://doi.org/10.1038/425037a.

24. Sadd, B.M.; Schmid-Hempel, P. Insect immunity shows specificity in protection upon secondary pathogen exposure. Curr Biol. 20, 1206-1210(2006). doi:10.1016/j.cub.2006.04.047.

25. Melillo, D.; Marino, R.; Italiani, P.; Boraschi, D. Innate Immune Memory in Invertebrate Metazoans: A Critical Appraisal. Front Immunol. 22, 1915 (2018). doi:10.3389/fimmu.2018.01915.

26. Green, T.J.; Montagnani, C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus (OsHV-1 μvar). Fish Shellfish Immunol. 35, 382–8 (2013). doi:10.1016/j.fsi.2013.04.051.

27. Lafont, M.; Petton, B.; Vergnes, A.; Pauletto, M.; Segarra, A.; Gourbal, B.; Montagnani, C. Long-lasting antiviral innate immune priming in the Lophotrochozoan Pacific oyster, *Crassostrea gigas*. Sci Rep. 7, 13143 (2017). doi:10.1038/s41598-017-13564-0.

28. Lafont, M.; Vergnes, A.; Vidal-Dupiol, J.; de Lorgeril, J.; Gueguen, Y.; Haffner, P.; Petton, B.; Chaparro, C.; Barrachina, C.; Destoumieux-Garzon, D.; Mitta, G.; Gourbal, B.; Montagnani, C. A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, *Crassostrea gigas*. mBio. 10, e02777-19 (2020). doi: 10.1128/mBio.02777-19.

29. Green, T.J.; Helbig, K.; Speck, P.; Raftos, D.A. Primed for success: oyster parents treated with poly(I:C) produce offspring with enhanced protection against Ostreid herpesvirus type I infection. Mol Immunol. 78, 113–20 (2016). doi:10.1016/j.molimm.2016.09.002.

30. Moreira, R.; Pereiro, P.; Canchaya, C.; Posada, D.; Figueras, A.; Novoa, B. RNA-Seq in *Mytilus galloprovincialis*: comparative transcriptomics and expression profiles among different tissues. BMC Genomics. 16, 728 (2015). doi:10.1186/s12864-015-1817-5.

31. Bjärnmark, N.A.; Yarra, T.; Churcher, A.M.; Felix, R.C.; Clark, M.S.; Power, D.M. Transcriptomics provides insight into *Mytilus galloprovincialis* (Mollusca: Bivalvia) mantle function and its role in biomineralisation. Mar genomics. 27, 37–45 (2016). doi:10.1016/j.margen.2016.03.004.

32. Moreira, R.; Pereiro, P.; Balseiro, P.; Milan, M.; Pauletto, M.; Bargelloni, L.; Novoa, B.; Figueras, A. Revealing *Mytilus galloprovincialis* transcriptomic profiles during ontogeny. Dev Comp Immunol. 84, 292–306 (2018). doi:10.1016/j.dci.2018.01.016.

33. Saco, A.; Rey-Campos, M.; Novoa, B.; Figueras, A. Transcriptomic response of mussel gills against *Vibrio splendidus* reveals its role in the immune response. Front Immunol. 11, 3273 (2020). doi:10.3389/fimmu.2020.615580.

34. Zhao, X.; Duan, X.; Wang, Z.; Zhang, W.; Li, Y.; Jin, C.; Xiong, J.; Li, C. Comparative transcriptome analysis of Sinonovacula constricta in gills and hepatopancreas in response to *Vibrio parahaemolyticus* infection. Fish Shellfish Immunol. 67, 523–535 (2017). doi:10.1016/j.fsi.2017.06.040.

35. Ren, Y.; Xue, J.; Yang, H.; Pan, B.; Bu, W. Transcriptome analysis of *Ruditapes philippinarum* hepatopancreas provides insights into immune signaling pathways under *Vibrio anguillarum* infection. Fish Shellfish Immunol. 64, 14–23 (2017). doi:10.1016/j.fsi.2017.03.005.

36. Collins, C.; Nelson, W.J. Running with neighbors: Coordinating cell migration and cell-cell adhesion. Curr Opin Cell Biol. 36, 62–70 (2015). doi:10.1016/j.ceb.2015.07.004.

37. Huse, M. Mechanical forces in the immune system. Nat Rev Immunol. 17, 679–690 (2017). doi:10.1038/nri.2017.74.

38. Bellavia, G.; Fasanaro, P.; Melchionna, R.; Capogrossi, M.C.; Napolitano, M. Transcriptional control of skin reepithelialization. J Dermatol Sci. 73, 3–9 (2014). doi:10.1016/j.jdermsci.2013.08.007.

39. Kasuya, A.; Tokura, Y. Attempts to accelerate wound healing. J Dermatol Sci. 76, 169–172 (2014). doi:10.1016/j.jdermsci.2014.11.001.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CONCLUSIONS

1. O transcriptoma de cada mexillón é completamente diferente, polo que o deseño dos experimentos debe ter en conta esta variabilidade individual.
2. Os mexillóns responden a sinais de perigo e lesións expresando péptidos antimicrobianos como a miticina C, que modula a resposta dos hemocitos ante infeccións posteriores.
3. Demostrouse unha reprogramación da expresión xénica do mexillón ante una situación de reinfección con *Vibrio splendidus*. Esta modificación da resposta inmune resultou no control do proceso inflamatorio, así como na orixe dun fenotipo de tolerancia inmunolóxica.
4. A miticina C provoca cambios no perfil de expresión xenética dos hemocitos, modificando o seu estado de activación e mobilidade.
5. En situacións de dano nos tecidos, a miticina C acelera os procesos de rexeneración tanto nun modelo humano de pel *in vitro* (HaCaT) como nun modelo de rexeneración de tecidos *in vivo* como o peixe cebra.
6. Identificáronse un total de 120 variantes de miticina diferentes, demostrando a gran diversidade deste xene. Estas variantes están suxeitas ao fenómeno de variación de presenza/ausencia.
7. A dispoñibilidade do xenoma e do transcriptoma do mesmo individuo confirmou que non hai procesos de edición de ARN detrás da variabilidade das miticinas e que a súa diversidade ten unha orixe xenómica.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA