Fish & Shellfish Immunology 35 (2013) 1899–1905

Contents lists available at ScienceDirect

Fish & Shellfish Immunology

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

Full length article

# Identification of immune-related SNPs in the transcriptome of *Mytilus chilensis* through high-throughput sequencing

Gustavo Núñez-Acuña, Cristian Gallardo-Escárate*

*Laboratory of Biotechnology and Aquatic Genomics, Interdisciplinary Center for Aquaculture Research (INCAR), University of Concepción, PO Box 160-C, Concepcion, Chile*

ARTICLE INFO

*Article history:*
Received 1 August 2013
Received in revised form
18 September 2013
Accepted 19 September 2013
Available online 28 September 2013

*Keywords:*
SNPs
*Mytilus chilensis*
Immunogenomics
RNA-seq

ABSTRACT

Single nucleotide polymorphisms (SNPs) identified in coding regions represent a useful tool for understanding the immune response against pathogens and stressful environmental conditions. In this study, a SNPs database was generated from transcripts involved in the innate immune response of the mussel *Mytilus chilensis*. The SNPs were identified through hemocytes transcriptome sequencing from 18 individuals, and SNPs mining was performed in 225,336 contigs, yielding 20,306 polymorphisms associated to immune-related genes. Classification of identified SNPs was based on different pathways of the immune response for *Mytilus* sp. A total of 28 SNPs were identified in the Toll-like receptor pathway and included 5 non-synonymous polymorphisms; 19 SNPs were identified in the apoptosis pathway and included 3 non-synonymous polymorphisms; 35 SNPs were identified in the Ubiquitin-mediated proteolysis pathway and included 4 non-synonymous variants; and 54 SNPs involved in other molecular functions related to the immune response, such as molecular chaperones, antimicrobial peptides, and genes that interacts with marine toxins were also identified. The molecular markers identified in this work could be useful for novel studies, such as those related to associations between high-resolution molecular markers and functional response to pathogen agents.



## 1. Introduction

One approach for contributing to the understanding of the immune system in aquatic species is the development of DNA markers such as SNPs (Single Nucleotide Polymorphisms), with a focus on the identification of different genotypes involved in the response to diseases and pathogens [1]. Various studies have emerged supporting the hypothesis that SNPs in immune-relevant genes could influence the susceptibility of individuals to pathogens and diseases. In aquatic species, polymorphisms relating to single genes, such as *inhibitors of serine proteases* in the Eastern oyster *Crassostrea virginica* [2] and *ubiquitin-conjugating enzyme E2* in the muricid *Concholepas concholepas* [3], have been described. In addition, SNPs identified from sequenced libraries have been reported in diverse marine species such as the Pacific oyster *Crassostrea gigas* [4], the turbot *Scophthalmus maximus* [5], and the common carp *Cyprinus carpio* [6]. In this latter species, SNPs were identified in a set of genes involved in the immune response, such as in *toll-like receptors*, *MyD88*, *TRAF*, *interferon*, and *interleukin-1β* [7]. Regarding specific genes, twelve SNPs were found in both the *TLR3* and *TLR22* genes of the grass carp *Ctenopharyngodon idella* [8–10]. In *Mytilus chilensis*, the target species of the current study, there are currently no studies regarding SNPs markers related to immune response genes.

Among the pathways involved in the immune response against different pathogens, the first barrier is comprised of diverse genes acting as Pattern Recognition Receptors (PRRs), which are generally located in the cellular membrane. Pathogen recognition is mediated by the interaction between PRRs and Pathogens-Associated Molecular Patterns (PAMPs), being the Toll-like Receptor (TLR) pathway the most studied route involved in this process [11–13]. In general, the TLR pathway is connected to the apoptosis pathway through the regulation of the *MyD88* and *caspase-8* genes. The apoptosis pathway triggers programmed cell death by regulating several genes, such as different isoforms of *caspases*, *TRAF*, *FADD*, and *MyD88*, among others. Thus, this pathway has a crucial impact on homeostasis [14]. Previous reports have found sequences of genes related to this pathway in shellfish, such as various *caspase* genes in *Mytilus galloprovincialis* [15] and two *BCL-2* genes in the Manila

* Corresponding author. Departamento de Oceanografía, Centro de Biotecnología, Universidad de Concepción, P.O. Box 160-C, Concepción, Chile.
*E-mail addresses:* crisgallardo@udec.cl, crisgallardo@oceanografia.udec.cl (C. Gallardo-Escárate).

1050-4648/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsi.2013.09.028

clam *Ruditapes philippinarum* [16]. Nevertheless, there are no SNPs identified. On the other hand, TLRs could be related to ubiquitin mediated proteolysis through the regulation of the NF-κB – IκBα complex. Ubiquitin-related genes have been characterized in diverse marine species, such as the *UBE2D* gene in the marine invertebrate, abalone *Haliotis diversicolor supertexta* [17]. In addition, a SNP was located in the *UBE2-like* gene in the marine gastropod *Concholepas concholepas*, and it was associated with genotypes that had different transcriptional response levels to an injection with *Vibrio anguillarum* [3]. Currently, this is the only work that has identified SNPs in genes involved in the ubiquitin-related proteolysis pathway in mollusks.

The innate immune system of mussels from the genus *Mytilus* has been studied under diverse approaches. In *M. galloprovincialis*, around 1820 sequences related to the immune response were collected in order to develop the first Immunochip for this genus [18]. In *Mytilus edulis*, a pyrosequencing approach was used to characterize a complex repertoire of immune-related genes concentrated in a certain pathway as a response against pathogens [19]. Regarding SNPs markers identified in these species, previous works have shown that *M. galloprovincialis* exhibits a high quantity of polymorphisms in the *myticin-C* gene [20], with 25 SNPs markers found for this species from public databases [21]. Moreover, 8 SNPs were used to study the genetic structure of three European species of the *Mytilus* taxa [22].

In regards to the target species of the current study, *M. chilensis*, the information is scarce about its immune response to pathogens. Until now, the NCBI Genbank database only lists 62 sequences for this species, and among these, about ten are related to genes with known functions. The expression levels for thirteen of these genes were triggered by a saxitoxin injection [23], but no SNPs have been currently identified in this species, despite its commercial and ecological relevance [24]. The objective of this study was to identify SNPs markers located in immune-relevant genes of the Chilean mussel *M. chilensis* in order to increase the genomic data and further immunological studies.

## 2. Materials and methods

### 2.1. Sample collection

Eighteen individuals were collected from Caleta Coliumo, Tomé, Bio Bio, Chile (36°32′S–72°57′W) and maintained at 14 °C in seawater with constant aeration. Around 1 mL of hemolymphs was collected from all individuals using sterile syringes, and then hemocytes were obtained through centrifugation in 1200 × *g* for 20 min at 4 °C. To achieve quality RNA extractions, hemocytes were pooled into groups of three mussels.

### 2.2. Total RNA extraction and transcriptome sequencing

Total RNA was extracted from each pool using the Ribopure™ kit (Ambion®, Life Technologies™, USA) following the manufacturer's instructions. Quantity, purity, and quality of isolated RNA were measured in TapeStation 2200 (Agilent Technologies Inc., Santa Clara, CA, USA) using the R6K reagents kit according to manufacturer instructions. Subsequently, double-stranded cDNA libraries were constructed using the TruSeq RNA Sample Preparation kit v2 (Illumina®, San Diego, CA, USA) and sequenced in the HiSeq 2000 platform (Illumina®) using 2x101 Pair-End read sequencing runs.

### 2.3. Bioinformatic analyses and SNPs discovery

Bioinformatic analyses were carried out using the CLC Genomics Workbench software (Version 6.0.1, CLC Bio, Denmark). De novo assembly was applied to mine SNPs variants from the transcriptome of six pools ($N = 3$ mussels) of *M. chilensis*. The assembly parameters used were the following: mismatch cost = 2, deletion cost = 3, insert cost = 3, minimum contig length = 200 bp, and trimming quality score = 0.05. Furthermore, to exclude paralogous sequence variants (PSVs) an overlap criterion of 70% and a similarity of 0.9 were applied [25]. Candidate SNPs were called with the following settings: window length = 11, maximum gap and

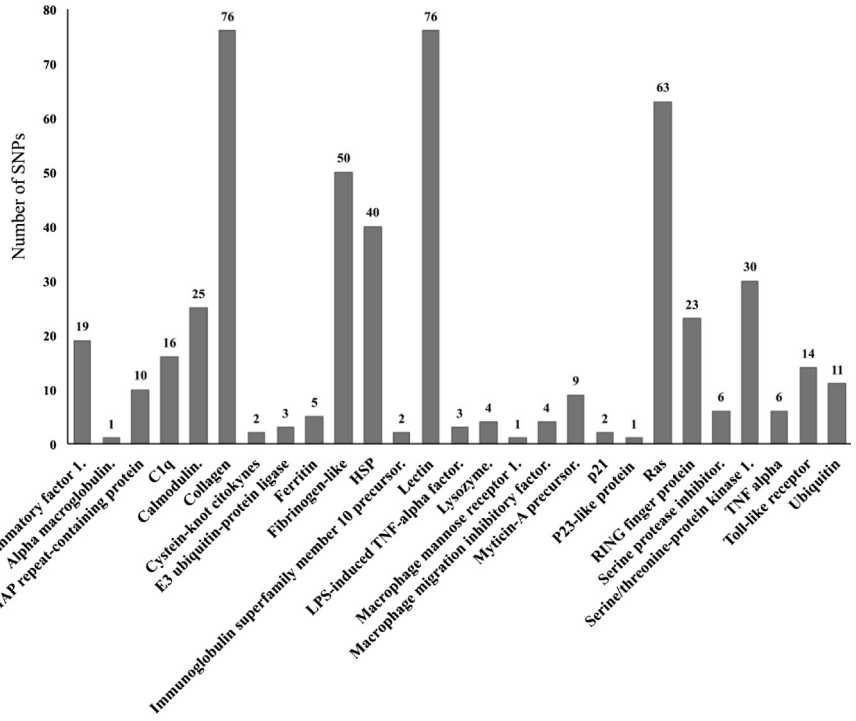


**Fig. 1.** Number of SNPs in immune-relevant genes of *Mytilus chilensis*. *Y*-axis correspond to count of the SNPs within the contigs annotated as the respective gene.

mismatch count = 2, minimum average quality of surrounding bases = 15, minimum quality of central base = 20, maximum coverage = 100, minimum coverage = 8, minimum variant frequency (%) = 35.0, and maximum expected variations (ploidy) = 2. In addition, a 454-homopolymer indels filter was applied.

Transcriptome annotation was based on a BLASTn analysis against a specific EST-database constructed with sequences of immune-relevant genes for mollusk species. The sequences were downloaded from the NCBI Genbank, Mytibase [26], and *M. edulis* EST-database [19]. Herein, sequences called "immune-relevant" were included in this specific database. In particular, this database was constructed downloading the sequences of all the genes in mammals and in fishes for six different KEGG pathways involved in the immune response: apoptosis, cytokine–cytokine receptor, JAK-STAT, p53, toll-like receptor and ubiquitin-mediated proteolysis pathways. Furthermore, Genbank sequences for mollusks were downloaded for other genes to be included in the database: allograft inflammatory factor, calmodulin, caspase, catalase, cathepsin, defensin, ependymin, ferritin, fibrinogen, glutathione-s transferase, heat shock-proteins, immunoglobulin, interleukin, lectin, lipopolysaccharide-induced TNF-alpha factor (LITAF), lysozyme, metalloproteinase, metallothionein, MyD88, myticin, mytilin, mytimacin, p53, peptidoglycan recognition protein (PGRP), plasminogen, super-oxide dismutase (SOD), serine protease inhibitors, toll-like receptor, TNF-alpha, ubiquitin. In total, the collection of all of these sequences encompassed a database of 5167 transcripts. Subsequently, SNPs variants were annotated according the BLAST hits. Further analyses, such as synonymous or non-synonymous SNPs identification, SNPs position in the mRNA (CDS, 3′UTR, 5′UTR), and classification by SNPs classes, were performed in the Geneious Software (version 6.0, Biomatter Ltda., New Zealand).

## 3. Results

### 3.1. SNPs in immune-relevant sequences: general overview

The transcriptome sequencing of *M. chilensis* yielded 799,899,594 reads which were assembled into 225,336 contigs with an average length of 372 bp. From this, 18,105 contigs (8.73%) were successfully annotated as immune-related genes through the BLAST analysis against the EST-immune-database. Furthermore, 8270 annotated contigs evidenced at least one SNP variant, and a total of 20,306 SNPs associated to immune-relevant sequences were counted (see details of SNP calling in Supplementary Table 1). As was expected, the majority of polymorphisms were first class SNPs (C/T and G/A), which numbered 10,845 and corresponded to 53.41%. Second class SNPs (C/A and G/T) numbered 4524 and

**Table 1**
List of most characteristic genes with SNPs in the different pathways analyzed.

| Gene | Function | Number of contigs | Number of SNPs | Best e-value | SNP position | | |
|---|---|---|---|---|---|---|---|
| | | | | | 5′ UTR | CDS | 3′ UTR |
| AKT | Toll-like receptor pathway | 1 | 1 | 5.70 E-160 | 0 | 1 | 0 |
| Calcium channel | Toxin binding | 4 | 11 | 5.53 E-84 | 2 | 3 | 6 |
| Calmodulin | Toxin binding | 6 | 10 | 3.98 E-110 | 1 | 7 | 2 |
| Calpain2 | Apoptosis pathway | 1 | 3 | 3.00 E-32 | 0 | 3 | 0 |
| CASP7 | Apoptosis pathway | 1 | 1 | 1.52 E-06 | 0 | 0 | 1 |
| Cullin1 | Ubiquitin mediated proteolysis pathway | 1 | 3 | 1.07 E-120 | 0 | 0 | 3 |
| Cullin3 | Ubiquitin mediated proteolysis pathway | 1 | 1 | 0 | 0 | 1 | 0 |
| Cullin4 | Ubiquitin mediated proteolysis pathway | 1 | 4 | 0 | 0 | 3 | 1 |
| CytC | Apoptosis pathway | 1 | 1 | 3.00 E-41 | 1 | 0 | 0 |
| G-type lysozyme | Antibacterial response | 1 | 1 | 6.56 E-22 | 0 | 1 | 0 |
| GABARAP | Ubiquitin mediated proteolysis pathway | 1 | 3 | 0 | 0 | 3 | 0 |
| HERC2 | Ubiquitin mediated proteolysis pathway | 1 | 5 | 8.14 E-150 | 1 | 4 | 0 |
| HSP22 | Molecular chaperone | 2 | 3 | 2.64 E-05 | 0 | 1 | 1 |
| HSP70 | Molecular chaperone | 5 | 13 | 0 | 0 | 9 | 4 |
| HSP90 | Molecular chaperone | 1 | 1 | 4.36 E-158 | 0 | 1 | 0 |
| IAP | Apoptosis pathway | 1 | 3 | 7.21 E-17 | 3 | 0 | 0 |
| IAP RP6 | Ubiquitin mediated proteolysis pathway | 1 | 1 | 2.30 E-52 | 0 | 1 | 0 |
| IFN-aBR | Toll-like receptor pathway | 1 | 1 | 1.20 E-06 | 0 | 1 | 0 |
| IkBa | Toll-like receptor pathway | 1 | 1 | 6.02 E-23 | 0 | 1 | 0 |
| IKKb | Toll-like receptor pathway | 2 | 8 | 7.23 E-83 | 0 | 8 | 0 |
| IRAK4 | Toll-like receptor pathway | 1 | 1 | 1.10 E-27 | 0 | 1 | 0 |
| IRF3 | Toll-like receptor pathway | 1 | 4 | 2.19 E-07 | 1 | 2 | 1 |
| L40 | Ubiquitin mediated proteolysis pathway | 1 | 1 | 5.33 E-58 | 0 | 0 | 1 |
| Lysozyme | Antibacterial response | 1 | 3 | 3.89 E-58 | 0 | 3 | 0 |
| MGD2b precursor | Antimicrobial peptide | 1 | 1 | 0 | 0 | 0 | 1 |
| MKK4 | Toll-like receptor pathway | 2 | 5 | 4.02 E-159 | 0 | 1 | 4 |
| Myticin-A precursor | Antimicrobial peptide | 1 | 9 | 1.33 E-100 | 0 | 0 | 9 |
| NF-kB | Toll-like receptor pathway | 1 | 1 | 2.07 E-86 | 0 | 1 | 0 |
| PfAF | Ubiquitin mediated proteolysis pathway | 2 | 3 | 2.62 E-107 | 0 | 3 | 0 |
| PKB | Apoptosis pathway | 5 | 11 | 1.28 E-85 | 0 | 11 | 0 |
| Polyubiquitin | Ubiquitin mediated proteolysis pathway | 1 | 2 | 4.43 E-60 | 0 | 2 | 0 |
| Potassium channel | Toxin binding | 1 | 1 | 8.77 E-89 | 0 | 1 | 0 |
| Prohibitin | Ubiquitin mediated proteolysis pathway | 1 | 1 | 4.68 E-77 | 0 | 1 | 0 |
| RING-box_protein_1 | Ubiquitin mediated proteolysis pathway | 1 | 1 | 6.93 E-90 | 0 | 0 | 1 |
| RIP1 | Toll-like receptor pathway | 1 | 1 | 2.37 E-33 | 0 | 1 | 0 |
| TLR1 | Toll-like receptor pathway | 1 | 1 | 9.88 E-27 | 0 | 1 | 0 |
| TLR3 | Toll-like receptor pathway | 3 | 3 | 1.44 E-14 | 0 | 1 | 0 |
| TRAF6 | Toll-like receptor pathway | 1 | 1 | 1.26 E-50 | 0 | 1 | 0 |
| Transferrin | Toxin binding | 1 | 1 | 7.21 E-16 | 0 | 1 | 0 |
| UBE2 | Ubiquitin mediated proteolysis pathway | 2 | 3 | 1.78 E-24 | 0 | 1 | 2 |
| UBE3 | Ubiquitin mediated proteolysis pathway | 2 | 2 | 6.17 E-30 | 0 | 1 | 1 |
| Ubiquitin_HIP2 | Ubiquitin mediated proteolysis pathway | 1 | 1 | 7.31 E-59 | 0 | 0 | 1 |

corresponded to 22.28%, while fourth class SNPs (A/T) numbered 4045 and corresponded to 4.39%, and finally, third class SNPs (C/G) numbered 892 and represented 4.39%. These classes correspond to the classification suggested by Liew et al. [27], which are associated to the differential melting temperature (Tm) among the complementary nucleotides.

The genes with the highest number of SNPs were also identified (Fig. 1). To conduct this analysis, sequences with the highest e-value (cutoff = 1E-30) score were selected and grouped into gene families when this corresponded. Interestingly, the maximum number of SNPs achieved in the same gene or gene family was obtained in *lectins* and *collagens*, which both had 76 polymorphisms. Following these genes, 63 polymorphisms were found in a homolog of the *Ras* gene. Other genes with a high number of SNPs were the *Fibrinogen-like* (50 SNPs), *heat shock protein* (40 SNPs), *serine/threonine protein kinase 1* (30 SNPs), *calmodulin* (25 SNPs), and *RING finger protein* (23 SNPs).

### 3.2. SNPs in the immune pathways of M. chilensis

Different genes that comprise the immune-related KEGG pathways were analyzed for specific SNPs (See summary in Table 1). The first pathway to be analyzed was the Toll-like receptors (TLR) signaling pathway. Among this set of genes, 28 SNPs were found. Of these, the kinase *IKKβ*, belonging to the *IKK* kinase complex, was the most polymorphic and had 30% of the SNPs. This gene was followed by the *Mitogen-activated protein kinase kinase 4* (*MKK4*) with 19% of the SNPs, and then the *Interferon regulatory factor 3* (*IRF3*) with 13% (Fig. 2A). The SNPs located in this pathway were mostly first class (50%) while the remaining classes composed the remaining quantity (Fig. 2B). SNPs were more frequently located in the coding regions (CDS) rather than in the untranslated regions (UTR) of these genes (Fig. 2C). From the total number of SNPs located in the coding region, there are four times more synonymous than non-synonymous substitutions (Fig. 2D). These non-synonymous SNPs correspond to the following four different substitutions: Glutamine to Histidine in the *IKKb* gene, Isoleucine to Methionine in the *IRAK4* gene, Leucine to Histidine in the *IRF3* gene, and Glycine to Valine in the *MKK4* gene.

Regarding the genes of the apoptosis pathway, 19 SNPs were found in five different genes. The *protein kinase B* gene (*AKT/PKB*) was the most polymorphic with 58% of the total number of SNPs, while the remaining SNPs were located in the *IAP*, *caspase 7*, *calpain 2*, and *CytC* genes (Fig. 3A). Class distributions, position of the SNPs in the genes, and amino acid substitution implications were proportionally similar to the TLR pathway (Figs. 3B–D). There were three non-synonymous substitutions that corresponded to a substitution between Tryptophan and a stop codon and Serine to Asparagine in the *AKT/PKB* gene, and Aspartic Acid to Glutamic Acid in the *calpain 2* gene.

On the other hand, genes involved in the ubiquitin-mediated proteolysis process were highly polymorphic and exhibited 35 SNPs. The distribution of the SNPs in this group were more homogenous than the other pathways, and the most polymorphic genes were in the *probable E3 ubiquitin-protein ligase HERC2* with 14% of SNPs, while *cullin 4* and *ubiquitin* had 11% of SNPs, and the remainder of SNPs were located in various genes with similar number of polymorphisms found in their sequences (Fig. 4A). Meanwhile, class distribution of SNPs, position, and the synonymous or non-synonymous ratio were similar to those in the other studied pathways (Fig. 4B–D). Non-synonymous substitutions corresponded to changes from Isoleucin to Asparagine in the *cullin 4* gene, Alanine to Aspartic Acid in the *IAP repeate-containing region 6* gene, Valine to Isoleucine in the *antisecretory factor-like protein*

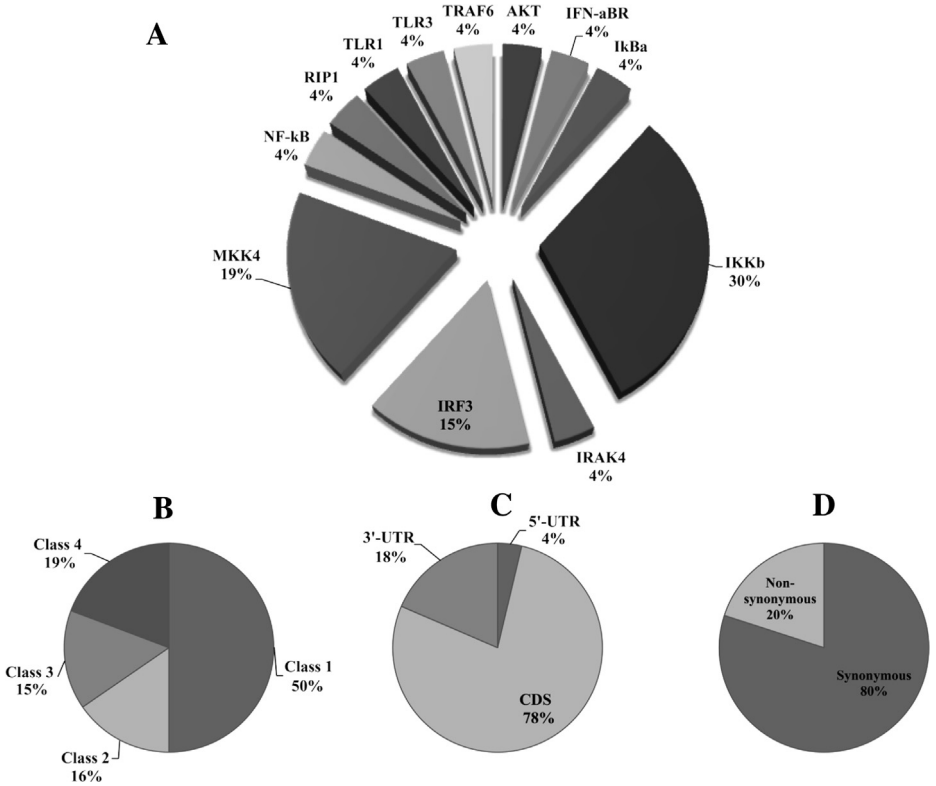


**Fig. 2.** Description of SNPs presented in the Toll-like receptor pathway of *Mytilus chilensis*. A: Distribution of SNPs by different genes of the pathway; B: Distribution of SNPs by classes; C: Distribution of SNPs according to its position in the transcript, D: Synonymous and Non-Synonymous SNPs proportion.

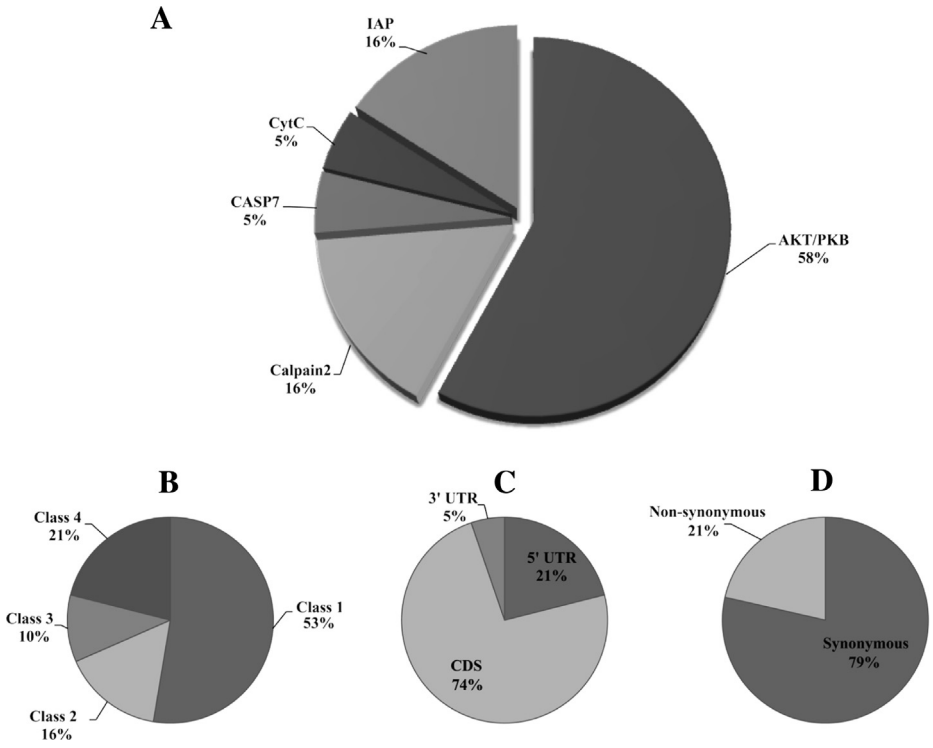


**Fig. 3.** Description of SNPs presented in the apoptosis pathway of *Mytilus chilensis*. A: Distribution of SNPs by different genes of the pathway; B: Distribution of SNPs by classes; C: Distribution of SNPs according to its position in the transcript, D: Synonymous and Non-Synonymous SNPs proportion.

*(PfAF)* gene, and Glycine to Cysteine in the *GABAA receptor-associated protein (GABARAP)* gene.

Finally, other genes involved in the response to environmental stress, toxins, and microbes were analyzed. From this set of genes we found high levels of polymorphisms in the molecular chaperone *HSP70*, in the *calcium channel* gene, in *Calmodulin*, and in the antimicrobial peptide *Myticin-A precursor*, among others (Fig. 5). The class distribution of these genes, position, and amino acid

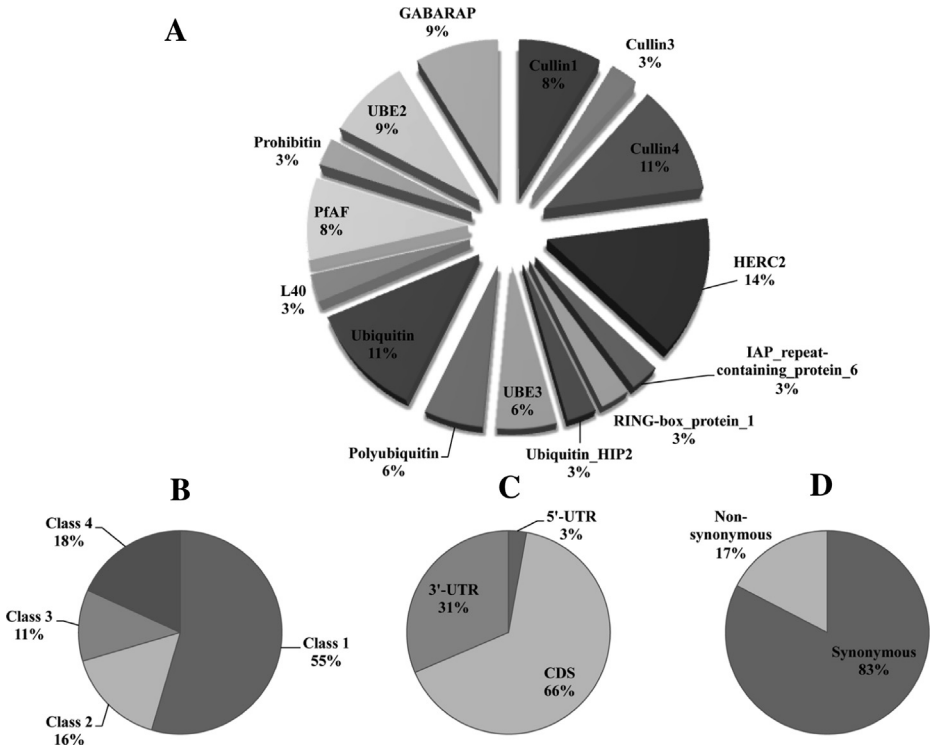


**Fig. 4.** Description of SNPs presented in the ubiquitin-mediated proteolysis pathway of *Mytilus chilensis*. A: Distribution of SNPs by different genes of the pathway; B: Distribution of SNPs by classes; C: Distribution of SNPs according to its position in the transcript, D: Synonymous and Non-Synonymous SNPs proportion.

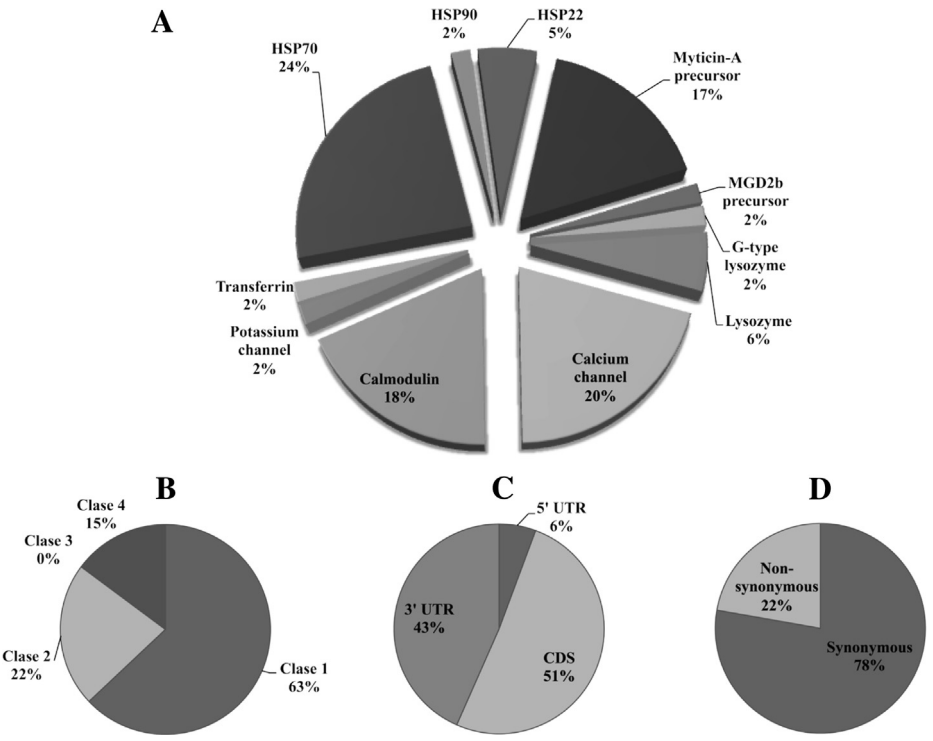


**Fig. 5.** Description of SNPs presented other genes involved in pathogen/stress response of *Mytilus chilensis*. A: Distribution of SNPs by different genes of the pathway; B: Distribution of SNPs by classes; C: Distribution of SNPs according to its position in the transcript, D: Synonymous and Non-Synonymous SNPs proportion.

changes were similar to the previously mentioned sets of genes (Fig. 5B–D). Some non-synonymous SNPs presented in these genes were as follows: Asparagine to Serine in the *HSP70* gene, Threonine to Alanine and Glutamine to Lysine in the *HSP22* gene, Aspartic acid to Glutamic Acid and Glycine to Aspartic acid in the *lysozyme* gene, and Alanine to Threonine and Methionine to Isoleucine in the *calmodulin* gene.

## 4. Discussion

The immune-related SNPs described in this work are a contribution to the immunogenomics field in *Mytilid* species. Currently, the 20,306 SNPs identified in this study are the widest set of markers involved in immune response for marine bivalves, and this amount only considers 8.73% of the total yield from the overall RNA-seq project that this study forms a part of. Undoubtedly, this RNA-seq project could be used in future works to span other biological processes of interest, which is more evidence of the strong potential of high-throughput sequencing to identify novel, high-resolution molecular markers [28].

Regarding in particular the molecular markers presented in immune-related genes, the SNPs found in this study have great potential for use in susceptibility/resistance evaluation analyses. Interestingly, in this study *lectins* and *collagens* were the genes that presented the highest number of SNPs. This is relevant, as previous studies have shown that SNPs in *collagenous lectins* are significantly associated with susceptibility to disease in pigs [29]. The SNPs analyzed in that study were present in both the collagenous-like and lectin domains. In the current study, high numbers of polymorphisms in relevant genes to these types of studies were found. For example, genes from the TLR pathway have been analyzed in association studies. The most polymorphic gene found in this pathway was *IKKβ*, which had only been associated with insulin resistance in humans [30] and obesity disease in porcine animals [31]. Another gene from this pathway with more than one polymorphism was *IRF3*, which has been proposed in genetic breeding programs in pigs due to its association with resistant traits [32]. On the other hand, previous studies associating genes from the apoptosis or ubiquitin-mediated proteolysis pathways with immune-related traits were not found, with the exception of the *UBE2-like* gene that was previously described by the authors of this study [3].

On the other hand, we found some genes with lower level of polymorphism than we expected according to previous descriptions. This is the case of some antimicrobial peptides such as different forms of the *Myticin* transcripts, which in other *Mytilus* species has shown a quite number of SNPs [20], while we found polymorphisms only in the *Myticin A* transcripts. Nonetheless, we could not establish a direct relationship because the previous works considered SNPs mining in both exons and intronic regions through sanger sequencing or pyrosequencing [33], while we considered only transcriptomic regions.

Despite the fact that previous works have developed SNPs markers in mussels [21,22,34], this is the first time that a wide transcriptomic sequence project was used to identify SNPs in immune-related genes in a species from the *Mytilus* genus. This survey to discover a wider number of SNPs in the immune-related transcripts in *M. chilensis* had the aim of developing useful markers for future studies, which could potentially associate different genotypes with susceptibility/resistance to pathogens and stress. Furthermore, SNPs markers in candidate genes represent a pivotal tool for understanding how the immune response could vary depending on the genotype of the specific locus.

## Acknowledgment

This study was funded by two grants provided by CONICYT-Chile (FONDECYT project 1120397 and FONDAP project 15110027).

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.fsi.2013.09.028.

## References


[1] Liu ZJ, Cordes JF. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 2004;238:1–37.
[2] Yu H, He Y, Wang X, Zhang Q, Bao Z, Guo X. Polymorphism in a serine protease inhibitor gene and its association with disease resistance in the eastern oyster (*Crassostrea virginica* Gmelin). Fish Shellfish Immunol 2011;30:757–62.
[3] Núñez-Acuña G, Aguilar-Espinoza A, Chávez-Mardones J, Gallardo-Escárate C. Ubiquitin-conjugating enzyme E2-like gene associated to pathogen response in *Concholepas concholepas*: SNP identification and transcription expression. Fish Shellfish Immunol 2012;33:1065–8.
[4] Zhang L, Guo X. Development and validation of single nucleotide polymorphism markers in the eastern oyster *Crassostrea virginica* Gmelin by mining ESTs and resequencing. Aquaculture 2010;302:124–9.
[5] Vera M, Álvarez-Dios JA, Millán A, Pardo BG, Bouza C, Hermida M, et al. Validation of single nucleotide polymorphism (SNP) markers from an immune Expressed Sequence Tag (EST) turbot, *Scophthalmus maximus*, database. Aquaculture 2011;313:31–41.
[6] Xu J, Ji P, Zhao Z, Zhang Y, Feng J, Wang J, et al. Genome-wide SNP discovery from transcriptome of four common carp strains. PLoS One 2012;7:e48140.
[7] Kongchum P, Rexroad CE, Hallerman EM, David L, Palti Y. Single nucleotide polymorphism identification, genetic mapping and tissue expression of the rainbow trout TLR9 gene. Anim Genet 2009;40:1001.
[8] Heng J, Su J, Huang T, Dong J, Chen L. The polymorphism and haplotype of TLR3 gene in grass carp (*Ctenopharyngodon idella*) and their associations with susceptibility/resistance to grass carp reovirus. Fish Shellfish Immunol 2011;30:45–50.
[9] Lv J, Huang R, Li H, Luo D, Liao L, Zhu Z, et al. Cloning and characterization of the grass carp (*Ctenopharyngodon idella*) Toll-like receptor 22 gene, a fish-specific gene. Fish Shellfish Immunol 2012;32:1022–31.
[10] Su J, Heng J, Huang T, Peng L, Yang C, Li Q. Identification, mRNA expression and genomic structure of TLR22 and its association with GCRV susceptibility/resistance in grass carp (*Ctenopharyngodon idella*). Dev Comp Immunol 2012;36:450–62.
[11] Kaisho T, Akira S. Toll-like receptor function and signaling. J Allergy Clin Immunol 2006;117:979–87.
[12] Kumar H, Kawai T, Akira S. Toll-like receptors and innate immunity. Biochem Biophys Res Commun 2009;388:621–5.
[13] Uematsu S, Akira S. The role of Toll-like receptors in immune disorders. Expert Opin Biol Ther 2006;6:203–14.
[14] Takle H, Andersen O. Caspases and apoptosis in fish. J Fish Biol 2007;71:326–49.
[15] Romero A, Estévez-Calvar N, Dios S, Figueras A, Novoa B. New insights into the apoptotic process in mollusks: characterization of caspase genes in *Mytilus galloprovincialis*. PLoS One 2011;6:e17003.
[16] Lee Y, Whang I, Lee S, Menike U, Oh C, Kang D-H, et al. Two molluscan BCL-2 family members from Manila clam, *Ruditapes philippinarum*: molecular characterization and immune responses. Fish Shellfish Immunol 2013;34:1628–34.
[17] Ye S, Wu L, Luo M, Xu T, Wu X. Characterization and function of an E2-17 kDa (UBE2D) in an invertebrate *Haliotis diversicolor supertexta*. Fish Shellfish Immunol 2013;34:1496–504.
[18] Venier P, Varotto L, Rosani U, Millino C, Celegato B, Bernante F, et al. Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics 2011;12:69.
[19] Philipp EER, Kraemer L, Melzner F, Poustka AJ, Thieme S, Findeisen U, et al. Massively parallel RNA sequencing Identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. PLoS One 2012;7:e33091.
[20] Pallavicini A, Costa MdM, Gestal C, Dreos R, Figueras A, Venier P, et al. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggest ancient host-pathogen interactions. Dev Comp Immunol 2008;32:213–26.
[21] Vera M, Pardo G, Pino-querido A, Álvarez-Dios JA, Fuentes J, Martínez P. Characterization of single-nucleotide polymorphism markers in the Mediterranean mussel, *Mytilus galloprovincialis* 2010:568–75.
[22] Zbawicka M, Drywa A, Śmietanka B, Wenne R. Identification and validation of novel SNP markers in European populations of marine *Mytilus mussels*. Mar Biol 2012;159:1347–62.
[23] Núñez-Acuña G, Aballay A, Hégaret H, Astuya A, Gallardo-Escárate C. Transcriptional responses of *Mytilus chilensis* exposed in vivo to saxitoxin (STX). J Molluscan Stud 2013:1–9. http://dx.doi.org/10.1093/mollus/eyt030 [in press].
[24] Núñez-Acuña G, Tapia FJ, Haye PA, Gallardo-Escárate C. Gene expression analysis in *Mytilus chilensis* populations reveals local patterns associated with ocean environmental conditions. J Exp Mar Biol Ecol 2012;420–421:56–64.
[25] Renaut S, Nolte AW, Bernatchez L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). Mol Ecol 2010;19:115–31.
[26] Venier P, De Pittà C, Bernante F, Varotto L, De Nardi B, Bovo G, et al. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 2009;10:72.
[27] Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, et al. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clin Chem 2004;50:1156–64.
[28] Aguilar Pierlé S, Dark M, Dahmen D, Palmer G, Brayton K. Comparative genomics and transcriptomics of trait-gene association. BMC Genomics 2012;13:699.
[29] Keirstead ND, Hayes Ma, Vandervoort GE, Brooks aS, Squires EJ, Lillie BN. Single nucleotide polymorphisms in collagenous lectins and other innate immune genes in pigs with common infectious diseases. Vet Immunol Immunopathol 2011;142:1–13.
[30] Fontanesi L, Galimberti G, Calo DG, Fronza R, Martelli PL, Scotti E, et al. Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian large white pigs using a selective genotyping approach. J Anim Sci 2012;90:2450–64.
[31] Miller MR, Zhang WM, Sibbel SP, Langefeld CD, Bowden DW, Haffner SM, et al. Variant in the 3′ region of the I kappa B alpha gene associated with insulin resistance in Hispanic Americans: the IRAS family study. Obesity 2010;18:555–62.
[32] Cai ZH, Wang LK, Liu DY, Li SH, Pang XS, Cheng YX. Polymorphism identification and association analysis of IRF3 gene in pig. J Anim Vet Adv 2011;10:1841–4.
[33] Rosani U, Varotto L, Rossi A, Roch P, Novoa B, Figueras A, et al. Massively parallel amplicon sequencing reveals isotype-specific variability of antimicrobial peptide transcripts in *Mytilus galloprovincialis*. PLoS One 2011;6:e26680.
[34] Tanguy A, Bierne N, Saavedra C, Pina B, Bachère E, Kube M, et al. Increasing genomic information in bivalves through new EST collections in four species: development of new genetic markers for environmental studies and genome evolution. Gene 2008;408:27–36.