REVIEW

How gene expression profiles disclose vital processes and immune responses in *Mytilus* spp.

S Domeneghetti[1*], C Manfrin[2*], L Varotto[1], U Rosani[1], M Gerdol[2], G De Moro[2], A Pallavicini[2], P Venier[1]

[1]*Department of Biology, University of Padua, Padua, Italy*
[2]*Department of Life Sciences, University of Trieste, Trieste, Italy*

*Equal contribution

Abstract

Gene expression studies largely support the understanding of gene-environment interactions in humans and other living organisms but the lack of genomic and genetic information often complicates the analysis of functional responses in non-traditional model species. Nevertheless, the fast advancement of DNA microarray and sequencing technologies now makes global gene expression analysis possible in virtually any species of interest. As regards the *Mytilus* genus, tens of thousands Expressed Sequence Tags (ESTs) are currently available for *M. californianus* and *M. galloprovincialis*, and DNA microarrays have been developed. Among them, Immunochip 1.0 specifically includes 1,820 probes of genes centrally involved or modulated in the innate immune responses of the Mediterranean mussel. This review recalls peculiarities and applications of the existing mussel DNA microarrays and finally summarizes facts concerning a variety of transcript sequences likely involved in the mussel immunity. Beside DNA microarrays, Next Generation Sequencing (NGS) technologies now offer new and broader research perspectives, from the whole transcriptome coverage to the *Mytilus* genome sequencing.

Key Words: *Mytilus*; DNA microarray; innate immunity; ESTs; antimicrobial peptides; C1q

Introduction

Global gene expression analyses in organisms selected to represent a given ecosystem currently support ecotoxicological investigations and create a conceptual bridge between the early organism responses and late population changes (Steinberg *et al.*, 2008). The animal response to a variety of detrimental conditions usually starts with alarm signals followed by adjustment reactions aimed to neutralize the physiological unbalance, and may end up in a general decline of vital processes ultimately marked by disease and death. Depending on the stress type and exposure intensity, the expression of definite sets of genes makes available specific proteins and other molecules in cells and tissues.

Appeared in the 1990s, the DNA microarray technology enables the simultaneous expression measure of thousands of genes represented in the microarray platform by unambiguous polynucleotide probes (Schena *et al.*, 1995; Lockhart *et al.*, 1996). The gene expression profiles emerging from suitable sampled cells or tissues can provide a dynamic view of biological processes and allow the correct sorting of different functional states. Based on the availability of sequence data, DNA microarrays can be used to solve a variety of biological questions: from the identification of molecular markers pathognomonic of disease and transcriptional signatures of various stress factors to the understanding of complex phenomena such as the epigenome in normality and disease (Martín-Subero and Esteller, 2011).

Specific microarray platforms and advanced deep sequencing technologies now support studies on the cellular functions of microRNAs and their role in human diseases (Thomas *et al.*, 2010). Leading research institutions are currently using both the mRNA and miRNA expression profiling to examine the genomic responses to environmental stresses (NCT). Central to the toxicogenomics studies is the concept of '*phenotypic anchoring*' which recalls the

Corresponding author:
Paola Venier
Department of Biology
University of Padua
Via U. Bassi, 58/B, 35121 Padua, Italy
E-mail: paola.venier@unipd.it
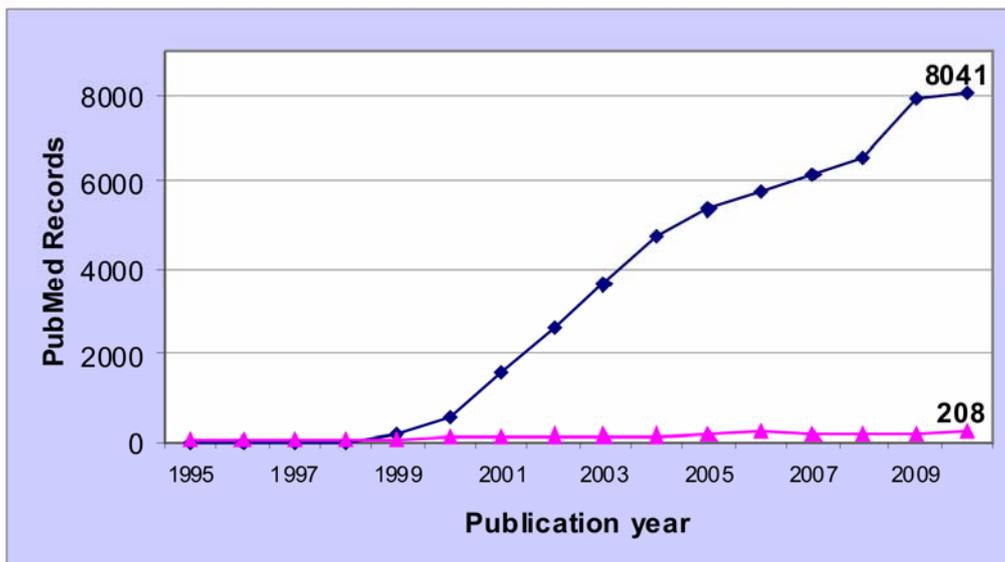
Fig. 1 Number of PubMed publications including the terms "DNA-microarray" (blue line) or "*Mytilus*" (purple line) from 1995 to 2010. Has the DNA microarray revolution reached its peak?

importance to correlate the observed gene expression changes to adverse effects defined by conventional parameters of toxicity and pathology.

In the controlled vocabulary of the Natl. Library of Medicine, the term 'DNA microarray' is indexed under the following category which indicates the large application range of such innovative technology (MESH): Oligonucleotide Array Sequence Analysis- the hybridization of a nucleic acid sample to a very large set of oligonucleotide probes, which are attached to a solid support, to determine sequence or to detect variations in a gene sequence or expression or for gene mapping.

Relevant to the gene expression profiling research area is Gene Expression Omnibus, a public repository that archives and freely distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomics data submitted by the scientific community (GEO). To fulfil the current standards (Minimum Information About a Microarray Experiment) the contents submitted to GEO should include the following: raw hybridization data; normalized data from which the main experimental findings can be outlined; description of the tested samples and whole experimental design, with details on the biological and technical replicates; identity and location of all probes and controls of the microarray platform, with external reference in the case of commercial arrays; concise but precise description of laboratory and data processing protocols related to the experiment under submission. According to the aims of the Microarray Gene Expression Data Society, dating back to the late '90s, the compliance to the MIAME standards should assure the data comparability among different platforms and testing protocols while supporting common work criteria and the reduction of random data variation (Rogers and

Cambrosio, 2007). Based on the comparative data analysis, the guidelines for standardization and reporting have been further refined (Chen *et al.*, 2007; Shi *et al.*, 2008). At present, GEO contains as much as 9,000 platform records which can be accessed and browsed in full detail.

Figure 1 illustrates the annual increase of PubMed records including the term "DNA-microarray" or "*Mytilus*" (subject heading or title/abstract) and suggests that pioneering technologies open the way to new ideas more than an unconventional model organisms. In fact, the gene expression profiling field has substantially diversified: specialized equipments and various related software make today the DNA microarrays powerful tools for the study of gene sequence, structure and expression, particularly for the best known model organisms. Nonetheless, one must remember that transcription is just one step in gene expression, and post transcriptional events referred to maturation of the primary transcript, RNA editing and RNA silencing as well as various modifications of the translation products overall influence the final amounts and activity of cellular proteins.

Mytilus DNA microarrays: preparation strategy and applications

Six GEO records refer to mussel DNA microarrays at July 2011.

MytArray 1.0 (GEO platform **GPL1799**, Oct 2006) is composed by 1,712 cDNA probes, univocally tagging the 3'-end region of transcripts from the main tissues of adult mussels (*Mytilus galloprovincialis*) and 46 unrelated cDNA control probes, all printed in duplicate and twice per slide (1.7 k mussel probes per array, 7.0 k total probes per slide). The probes were designed in the 3'-UTR,
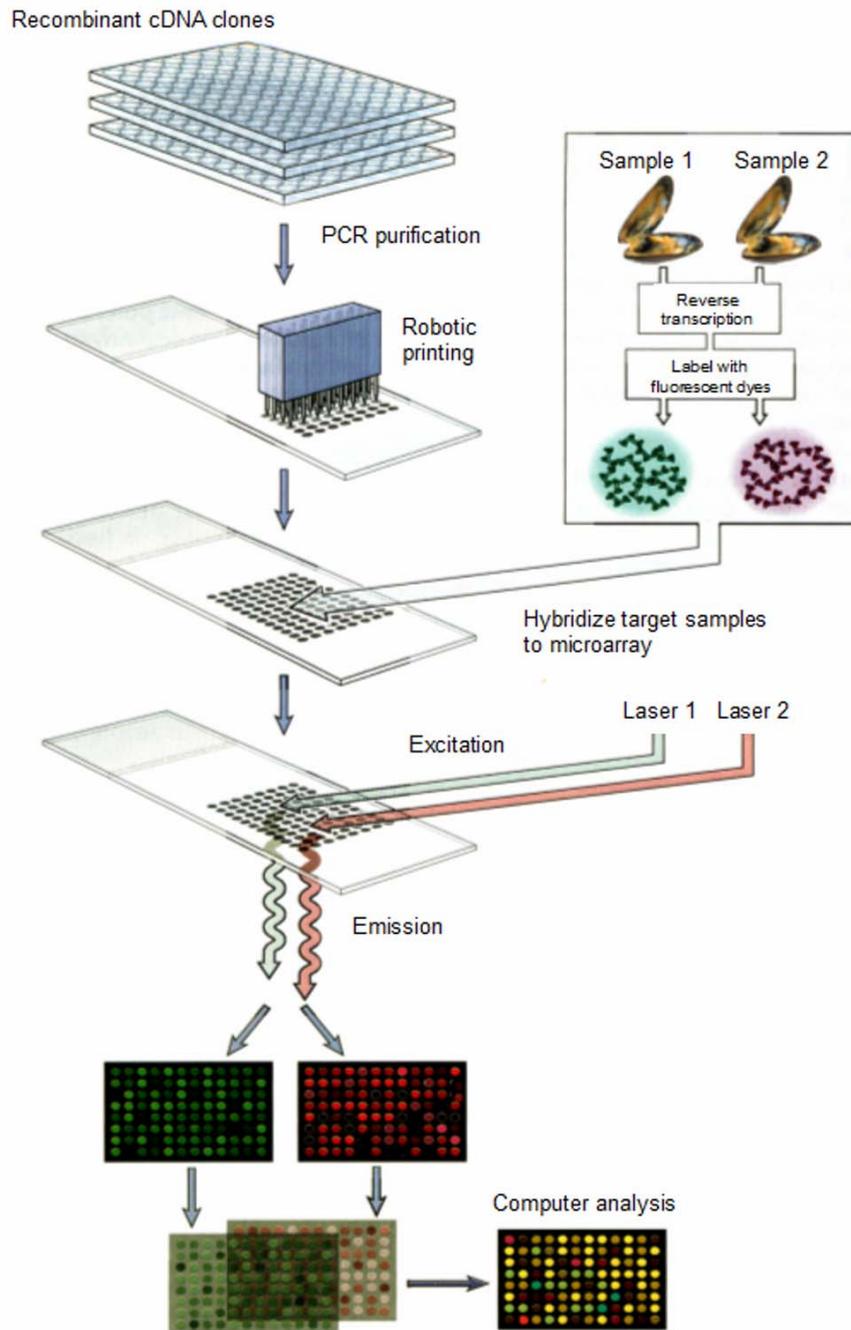
Fig. 2 Work diagram referred to the competitive hybridization of two dye-swap-labelled samples on a cDNA microarray with two-channel detection of the fluorescence signals (modified from Gibson and Muse, 2004).

one among the least conserved gene regions, so that competition of different mRNAs from genes with similar coding sequence and cross-hybridization to the same microarray probe should be minimal. Also, the probe size of 400 - 800 bp is expected to ensure comparable efficiency in the amplification and spotting of the cDNA inserts as well as uniform hybridization kinetics (Venier et al., 2006).

MytArray 1.0 was first used to investigate the specificity of gene transcription in mussel tissues with different functional role and the transcriptional profiles of mussels treated with chemical mixtures or living wild in different sites of the Venice lagoon (Venier *et al.*, 2006; GEO series **GSE2176**, **GSE2183** and **GSE2184**). Sample pairs combined according to dye-swap labelling (reference and test samples labelled with Cy3/Cy5 cyanine dyes in alternate combinations) were competitively hybridized on the two equal arrays of cDNAs spotted on the same slide (Fig. 2). Gills, digestive

gland, tissues involved in contraction/motility (foot, adductor muscles, ligaments) and reproduction (gonads and mantle) displayed specific transcriptional footprints, as expected. The results obtained in mussels treated with mixtures of inorganic metal salts or persistent organic chemicals guided the interpretation of the gene expression profiles of mussels living in the inner industrial canals or at the lagoon border open to the sea (this exercise yielded a provisional list of contamination marker probes). In this study, the evident transcriptional down-regulation detected in the reproductive tissues was consistent with the depleted status of the mussel gonads whereas the greatest variety and abundance of transcripts was found in the digestive gland. Additional analysis of these expression data is reported elsewhere (Pantzartzi *et al.*, 2010).

The same platform was then used to evaluate in a time-course study the gene expression changes in the digestive gland of mussels exposed to okadaic acid (OA) via food contamination for five weeks (Manfrin *et al.*, 2010; GEO series **GSE14885**). One relevant purpose of the study was the identification of molecular biomarkers which could enable an easy and rapid detection of the Diarrhoeic Shellfish Poisoning biotoxins in marketable mussel stocks, *i.e.*, novel reliable assays complementing the existing diagnostic methods. An unsaturated loop design, combining control and treated samples with different dye-labelling for the competitive hybridization on Mytarray 1.0, was adopted to take into account all the time points and the biological replicates, with some combinations only inferred (Kerr and Churchill, 2001). A considerable number of transcriptional changes was detected in the OA-exposed mussels, with a prevalence of up-regulated probes at 3 days and a subsequent progressive increase of down-regulated probes (from 58 % over-expressed to 76 % under-expressed genes, respectively detected at day 3 and day 35). The biphasic time-related trend of response observed in this study recalls the changes occurring in the mussel digestive gland along different phases of the mussel reaction to the experimental stimulus, from the early acute response to the late overall unbalance of the functional processes. Many candidate markers are now under study to evaluate their predictive value in the diagnosis of biotoxin-contaminated mussels.

MytArray 1.1 (**GPL102699**, March 2010) contains the same cDNA probes of MytArray 1.0 in a slightly modified platform geometry. It has been used to study the gene expression profiles of *M. galloprovincialis* with monthly samplings for one year, hence taking into account seasonal differences which are known to influence metabolism rates and gonad development among other vital functions (Banni *et al.*, 2011; GEO series **GSE22915**, **GSE23049**- **GSE23051**). Mussels were collected from an anthropized and industrialized lagoon of the Southern Mediterranean Sea (Ben Said *et al.*, 2009) and competitive hybridizations were performed with dye-swap-labelled samples (dual colour analysis). Following a loop design with 3-4 biological replicates and parallel histological evaluation of the gonad status, the authors could analyze the transcriptional profiles of digestive gland tissue of female mussels collected during 12 months, and those of digestive gland and mantle tissues from male and female individuals representing all four gonad maturation stages. In the examined annual period, the transcriptional profiles globally highlighted the higher expression of genes associated to mussel nutrition and digestion in May-August compared to the other months, and trends for gonad transcripts consistent with the reproductive mussel status.

The same cDNA platform contributed to the toxicological evaluation of a neonicotinoid insecticide mixture (Dondero *et al.*, 2010), an organophosphate compound (Canesi *et al.*, 2011) and to the integrated measure of the functional mussel responses in the estuarine Tamar region in UK (Shaw *et al.*, 2011).

The Hofmann_UCSB_Mytilus_2.5K_v1.0 record (**GPL5795**, Mar 2008) describes a platform of nearly 2500 spotted cDNAs of *Mytilus californianus* consisting of both unsequenced and sequenced clones referring to gill and muscle of environmentally challenged mussels. The related GEO series **GSE8935** include data on latitudinal gene expression changes. Five biological replicates from four populations of Californian mussels were compared to a common reference sample in dual colour analysis (dye-swap labelling).

The HMS/SomeroLab-Mytilus-105K array-v1.0 (**GPL9676**, Jun 2010) and HMS/Somero-Mytilus-105K Agilent-v1.0 salinity stress (**GPL11156**, Jan 2011) are two successive versions of a platform composed by oligomer probes *in-situ* synthesized by Agilent Technologies (Santa Clara, CA, USA). These microarrays include probes of both *M. californianus* and *M. galloprovincialis,* and are intended for homologous and heterologous gene expression profiling. The processing and assembling of about 26,000 ESTs from *M. californianus* (Gracey *et al.*, 2008) and 3,984 ESTs from *M. galloprovincialis* (Venier *et al.*, 2003) resulted in a total of 12,961 and 1,688 transcript clusters or singletons, respectively. Long (60-mer) oligoprobes were designed against the *M. californianus* series and the resulting 43,969 total unique probes (2.6 probes per transcript sequence) were analyzed through BLAST searches against the *M. galloprovincialis* series to support selection and design of related probes (556 probe pairs matching transcripts of both species, with a mean number of 4.6 divergent nucleotide bases per probe). A total of 44,524 unique probes were duplicated or triplicated randomly to fill a microarray of 105,000 elements (105 k probes).

These two platforms have been used to investigate the transcriptional responses to thermal and osmotic stresses in *M. californianus*, *M. trossulus* and *M. galloprovincialis* (Evans and Somero, 2010; Lockwood *et al.*, 2010; Lockwood and Somero, 2011). To control the effects of sequence mismatches in the case of *M. galloprovincialis* probes included in the GPL9676 platform, only probes experimentally confirmed in the hybridization of 84 samples of both *M. galloprovincialis* and *M. trossulus* were used in the related data analysis. Following a large set of

hybridization experiments and stringent quality control, misleading probes were removed from the dataset and the second platform version (GPL11156/Agilent 019153) was generated.

In the central and southern coasts of California, *M. galloprovincialis* has largely displaced the native congener, *M. trossulus*, and such evidence could be explained by species differences in physiological traits related to the adaptation to warm habitats. To investigate the hypothesis, gene expression profiling was performed on gill RNA from mussels subjected to acute heat-stress (GEO series **GSE19031**). A total of 1,531 probes, out of 4,488 different genes represented on the microarray and recognizing mRNAs of both species, showed temperature-dependent expression changes highly similar in the two congeners whereas 96 probes denoting oxidative stress, proteolysis, energy metabolism, ion transport, cell signalling, and cytoskeleton reorganization outlined species-specific responses to the heat-stress. Among them, the one encoding the small heat shock protein 24 was highly induced in the Mediterranean mussel and showed only a small change in *M. trossulus*. Six biological replicates per mussel group were included in this study which exemplifies the use of a cross-species microarray as well as heterologous and homologous hybridization. According to the authors and published literature*, M. trossulus* and *M. galloprovincialis* are approximately 7.6 million years divergent from *M. californianus*, and only 3.5 million years divergent from each other: in other words, the heterologous hybridization of target sequences from *M. trossulus* should occur on microarray probes from *M. galloprovincialis* without inherent sequence bias and should provide a reliable comparison of their transcriptional responses. Though debated, prudent evaluations of the sequence divergence by *in silico* approaches and phylogenetic data could expand the use of cross-species hybridization as a compromise solution for investigating gene expression in species with unsequenced genomes (Costa *et al.*, 2010; Nazar *et al.*, 2010; Ptitsyn *et al.*, 2010).

Gene expression profiling was also performed on gill RNA from mussels subjected to salinity stress (GEO series **GSE25111**). A total of 117 probes, out of 6,777 genes represented on the microarray, showed significant changes similar between *M. californianus* and *M. galloprovincialis* whereas 12 probes, denoting mRNA splicing, polyamine synthesis, exocytosis, translation, cell adhesion, and cell signaling, outlined species-specific responses. The study was based on AlexaFluor-labelling (555 and 647 fluorescence dyes) of amplified RNA, pooled reference samples, six biological replicates, and competitive hybridization in agreement to the recommended Agilent protocols. In addition to the overall stringent processing of the fluorescence signals, the heterologous hybridization design suggested the elimination of data from probes with low signal intensity (signal intensity < 150 % of the local background and hybridized spot diameter < 30 % of the nominal spot diameter).

The work performed at the A. Gracey's and G.N. Somero's laboratories (University of Southern California -Los Angeles, CA, U.S.A. and Stanford University -Palo Alto, CA, U.S.A., respectively) on *Mytilus* (GEO series GSE19031 and GSE25111) and other species is facing the fundamental aspects of the organism adaptation to fluctuating environments and global climate changes, and gene expression profiling has been essential to their findings. For instance, the study of gene-expression changes in the Californian mussels at different phases in the tidal cycle revealed at least four distinct physiological states, corresponding to metabolism and respiration phase, cell-division phase, and two stress-response signatures linked to moderate and severe heat-stress events. The metabolism and cell-division phases appeared to be functionally linked and anti-correlated in time whereas magnitude and timing of the above states resulted to be influenced by the microhabitat conditions according to the vertical position on the shore (Gracey *et al.*, 2008). Based on comparative physiology, a recent paper offers an overview on the expected consequences of global climate changes (Somero, 2011).

Finally, the Mussel Immunochip 1.0 (**GPL10758**, April 2011) is a spotted oligonucleotide platform consisting of four-replicated 1820 oligomer probes plus unrelated controls prepared at CRIBI for the purposes of a recent European project (IMAQUANIM). Oligomers of 57 bases average length were designed at short distance from the 3' end of transcript sequences selected previously in Mytibase, the interactive knowledgebase of *M. galloprovincialis* which includes most of the ESTs publicly available for this species (Venier *et al.*, 2009). Based on multiple criteria, the subset of transcripts selected from Mytibase as putatively immune-related molecules should denote central "players" of the mussel innate immunity or genes whose expression is modulated during the mussel responses to immunostimulation (Venier *et al.*, 2011). In the platform description, the probe ID is hyperlinked to the relative Mytibase record: for instance the probe MGO_07346 relates to MGC07346, a mussel transcript featured by the protein domain IPR000098-Interleukin 10 and yet functionally unknown. The performance of Immunochip 1.0 was tested with hemolymph samples collected at 3 and 48 h from *Vibrio*-challenged mussels (GEO series **GSE23535**) according to competitive hybridization of dye-swap labelled amplified RNA samples.

In agreement with the above descriptions, Figure 3 provides an updated summary of the nucleotide and protein sequences publicly available at July 2011 and highlights the importance of EST sequencing for the preparation of new DNA microarrays. More about the molecular "players" of the innate immunity and the immune responses of *M. galloprovincialis* is reported in the following paragraph.

How much can simple sequences tell us about the mussel immune responses?

Taking advantage of the continuous increase of the nucleotide and amino acid sequences in the public databases, the current methods of bioinformatics can extract instructive data from
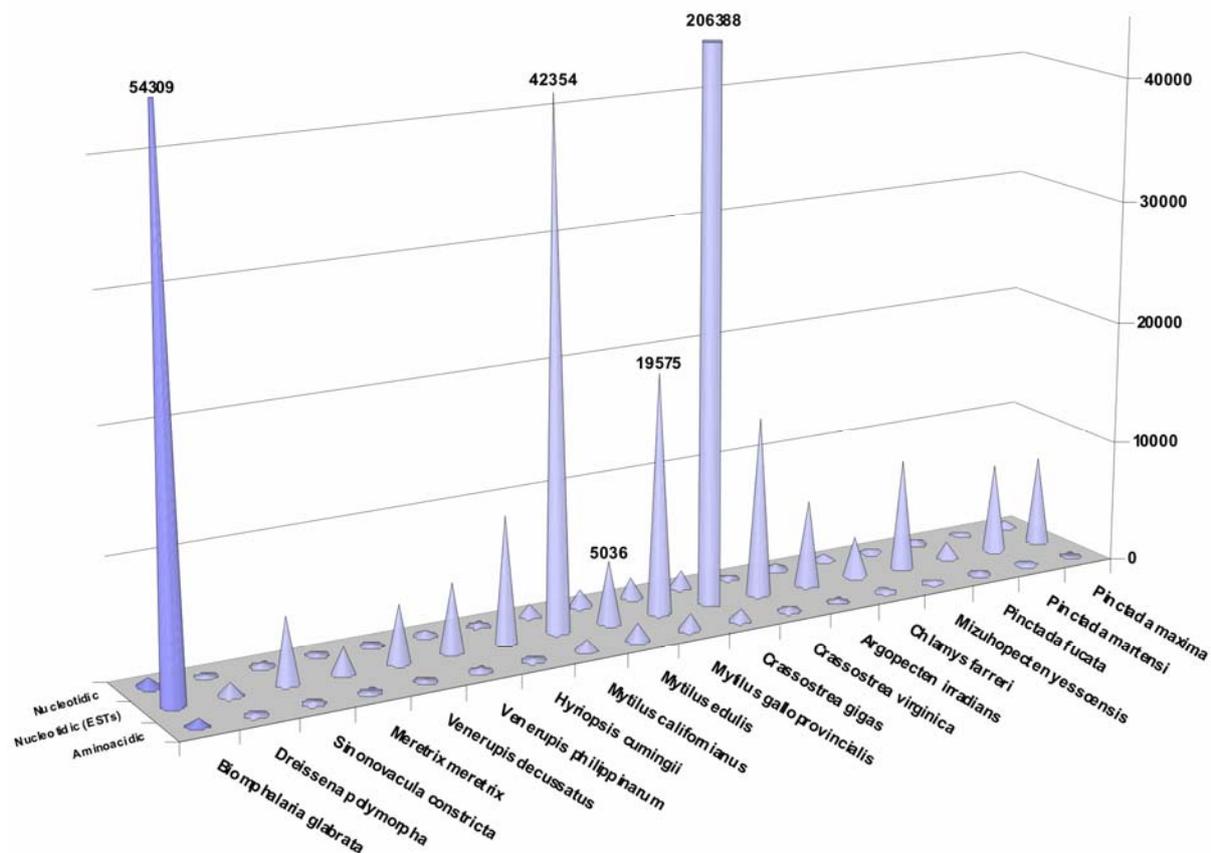
Fig. 3 Number of sequence records available for selected mollusc species at the Natl Center for Biotechnology Information at July 2011. DNA, RNA and protein sequences refer to *Biomphalaria glabrata* (Gastropoda) and bivalves belonging to the Veneroida, Unionoida, Mytiloida, Ostreoida, Pectinoida and Pterioida orders.

simple sequences: from the analysis of various gene/transcript regions to the evaluation of protein/peptide structure and to the comparative analysis of evolutionary differences across the tree of life. This procedural approach complements and integrates the data derived from long-standing disciplines such as measures of structural changes and protein amounts/activity, among others.

The overall analysis of 18,788 high-quality ESTs rationally organized in 7,112 independent clusters or singletons (Mytibase transcript collection) highlighted some particularly abundant transcript groups: namely, transcripts featured by a complement component C1q-like domain, antimicrobial peptide (AMP) precursors of all four families known in the Mediterranean mussel and many heterogeneous lectins including fibrinogen-related molecules (Venier *et al.*, 2011). To explain the abundance of immune-related molecules in Mytibase it is important to remember that such collection has been prepared by 16 primary (5 from hemocytes) and 1 normalized cDNA libraries from mussels subjected to various challenges, for instance mussels immune stimulated with preparations of Gram positive and Gram negative cells and viral-like molecules.

Searches by protein domain revealed a total of 168 different Mytibase transcripts containing the

C1q signature IPR001073, almost invariably associated with the overlapping TNF-like IPR008983 motif. Curiously, the C1q domain-containing proteins predicted from the transcript sequences, display a short N-terminal signal peptide and a C-terminal globular domain but no central collagen-like repeats which are instead typical of vertebrate C1q domain-containing proteins. According to the current literature, these mussel proteins could represent secreted globular receptors, components of ancient complement pathways expected to mediate pathogen recognition and lysis (Dodds and Matsushita, 2007). The modularity and versatility of binding mediated by the globular C1q domain explain the variety of roles currently attributed to this still expanding family of proteins, and also supports their involvement in pathogen pattern recognition (Carland and Gerwick, 2010). The abundance and variety of mussel C1q domain-containing transcripts are consistent with this view.

One among these transcripts, named MgC1q, resulted to be expressed at detectable levels in the main tissues of naïve adult mussels, with the hemocytes showing the highest expression levels, and from 2 h post-fertilization up to 3 months later. The MgC1q expression was significantly modulated after mussel infection with Gram positive or Gram

negative bacteria, data which confirm MgC1q as an immune-related gene. The striking molecular diversity of MgC1q was confirmed at both the DNA and cDNA levels, hence posing mechanistic questions on the origin of such variation (Gestal *et al.*, 2010). Experimental findings and sequence analyses support the hypothesis of gene duplication, functional diversification and positive selection of many C1qDC variants in selected taxa, including the mussel lineage (Gerdol *et al.*, 2011).

Defensins, mytilins, myticins and mytimycins are cationic antimicrobial peptides stabilized by 4 intrachain disulphide bonds (6 in mytimycin) in a typical 3-D motif (Yeaman and Yount, 2007). A remarkable diversity of a new group of myticins, with specific variant profiles detectable in single mussels, was reported in *M. galloprovincialis* (Pallavicini *et al.*, 2008; Costa *et al.*, 2009). Following the discovery of the myticin-C variants, their molecular diversity and evolution has been further discussed (Padhi and Verghese, 2008) and the most recent findings indicate myticin C as a chemotactic molecule with antiviral activity and immunoregulatory properties (Balseiro *et al.*, 2011). Just one singleton and other four similar sequences denote the antifungal AMP mytimycin in Mytibase (rare transcript). Mytimycin is composed by 54 aminoacids (6.2 - 6.3 kDa, 12 cysteines) and two main precursor variants, both featured by a signal peptide and a C-terminal extension, are expressed in mussels from different European regions (Sonthi et al., 2011). The presence of a calcium binding (EF hand) motif in the C-terminal extension suggests further characterization of such unusual AMP.

The "effector" role of the mussel antimicrobial peptides (AMPs) is confirmed in many experimental studies and a comprehensive review have been recently provided (Li *et al.*, 2011). Whether these effectors can modulate the mussel immune responses with mechanisms other than membrane disruption, as reported for mammalian AMPs, it is not clear. Based on deep amplicon sequencing, the sequence diversity of mussel AMPs is now under study in natural mussel populations from different geographical regions and in mussels challenged with bacterial cells.

Lectins are a rather heterogeneous protein family comprising 8 to 15 subgroups, depending on the scientist's view (Dodd and Drickamer, 2001). Lectins typically possess carbohydrate binding domains and participate in many cell processes. Similarly to the mammalian C1q, the C-terminal fibrinogen-like domain IPR002181 of ficolins forms a tulip-like structure able to bind the carbohydrate residues of foreign and apoptotic cells (with consequent opsonization, phagocytosis and cell clearance) or triggering the proteolytic complement cascade and pathogen lysis. Fibrinogen-related lectin proteins (FREPs) are expressed also in mussels (Venier *et al.*, 2011) and are codified by at least 2 (*M. edulis*) 4 (*M. californianus*) and 7 genes (*M. galloprovincialis)* (Gorbushin and Iakovleva, 2011). These molecules can be regarded as immune pattern-recognition receptors and their involvement in the native immunity is supported by the evidence of species-specific expansion of FREPs in the snail *Biomphalaria glabrata* and the mosquito *Anopheles gambiae* (Waterhouse *et al.*, 2007; Zhang *et al.*, 2008). In mussel, FREPs are significantly up-regulated after bacterial infection or PAMP treatment, and display opsonizing activity similar to that of mammalian ficolins; moreover, the different sets of FREP sequences detected among and within individuals further emphasize the great complexity of the invertebrate immune systems (Romero *et al.*, 2011). Other lectin-like sequences expressed in mussels are commented in Venier *et al.* (2011).

The cases reported above are a few examples of the many classes of transcripts specifically expressed or modulated during the mussel response to potential pathogens. Considering in a dynamic view the behaviour of one cell population only, the versatile mussel hemocytes, one can imagine that almost all cellular processes could be influenced by the contact with pathogen-associated molecular patterns: from the cytoskeleton remodelling supportive of chemotaxis, migration and phagocytosis to the intracellular signalling possibly shaping the inflammatory response and finely tuned expression of many regulatory and effector genes. Cross-talking signalling pathways have been traced in mussel and the Mytibase collection includes transcripts denoting the regulatory cytokine MIF (migration inhibiting factor) and cytokine-related molecules, consistent with the idea of an invertebrate cytokine network (Malagoli, 2010). The recent definition of a species-specific Immunochip aims to the experimental validation of a selected subset of transcripts: a synopsis of the main gene expression changes detected in mussels at 3 and 48 h after challenge with live bacterial cells is reported in Fig. 4. The general AMP down-regulation observed in this particular laboratory treatment was confirmed by quantitative PCR data and is discussed also in Li *et al.* (2010).

Concluding remarks

EST sequencing and DNA microarrays have substantially improved the identification of genes expressed in the *Mytilus* species. Compared to the first EST collection and the related cDNA microarray, Mytibase includes an interesting variety of immune-related molecules which can be further characterized with traditional and innovative approaches as exemplified by Romero *et al.* (2011). Nonetheless, in the Mytibase collection about half of the mussel transcripts are still unknown, devoid of functional annotation. Hence, much work remains to be done both *in silico* and in laboratory to provide a comprehensive view of the global gene transcription in mussels, particularly the part of the transcriptome mediating the response to potential invaders (immunome).

Undoubtedly, the application of the available mussel DNA microarray platforms can further reveal expression trends of different gene categories and identify useful markers of functional state, if not global molecular signatures useful to disentangle the complex mussel physiology. Depending on the study design and on the type of microarray platform, independent validation of the expression data can be accomplished by quantitative PCR or with other
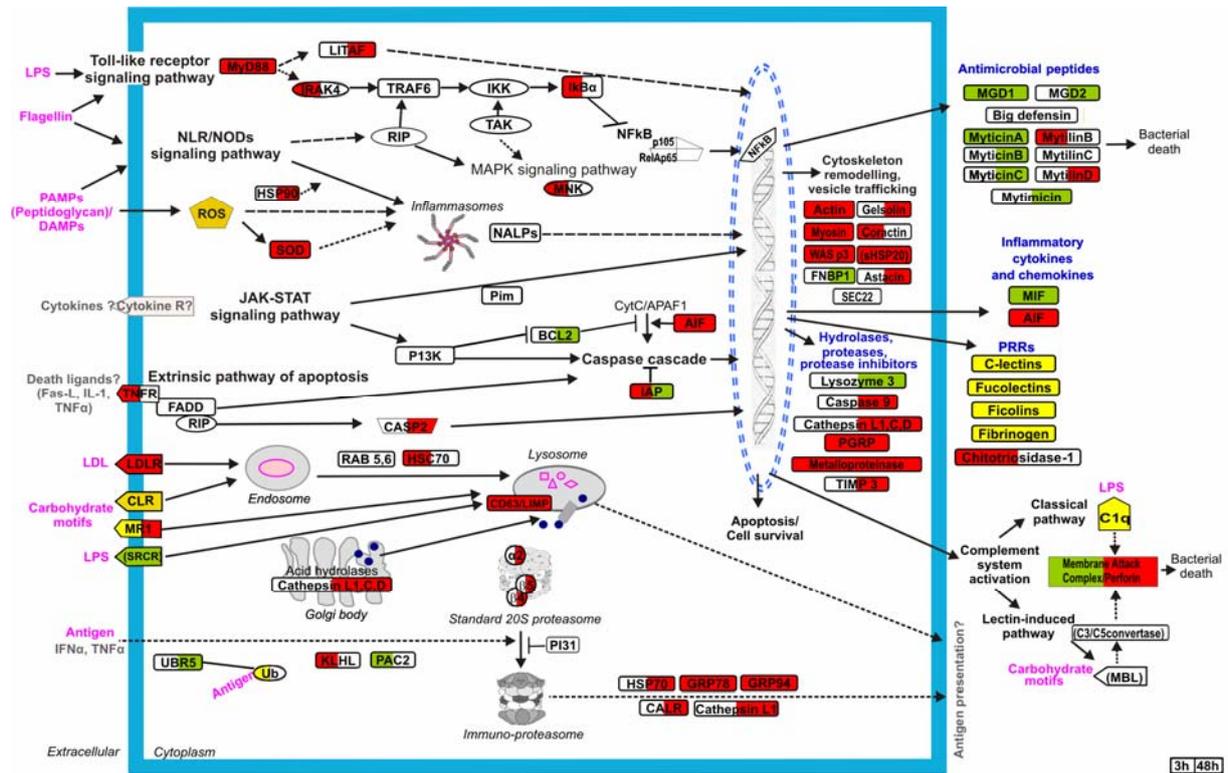
Fig. 4 Main transcriptional changes detected in mussels at 3 h and 48 h from the injection of live *Vibrio* cells (modified from Venier *et al.*, 2011). Only relevant molecular "players" represented in the Immunochip of *M. galloprovincialis* are reported (framed). In each frame, the detected expression trends are indicated in red, green and yellow (up- and down-regulation and not homogeneous trends, respectively). Annotations based only on protein domains are reported in brackets. Overall, the figure draws the attention to a number of mussel genes, still not characterized, whose expression is modulated in response to immune stimulation.

Abbreviation list (Fig. 4):

AIF: Allograft Inflammatory Factor
APAF1: Apoptotic Peptidase Activating Factor 1
BCL2: Baculoviral apoptosis regulator 2
C1-C5: Complement component 1-5
CALR: Calreticulin
CASP: Caspase
CD63/LIMP: Tetraspanin-7 (lysosome membrane protein)
CLR: C-type Lectin Receptor
CLR: C-type Lectin Receptor
DAMPs: damage-associated molecular patterns
FADD: FAS (TNFRSF)-Associated via Death Domain
FNBP1: Formin-Binding Protein 1
GRP 78/94: Glucose-Regulated Protein 78/94
HSC70: Heat Shock Cognate 70
HSP70/90: Heast Shock Protein 70/90
IAP: Inhibitor of Apoptosis Proteins;
IKBα: Inhibitor of nuclear factor Kappa-B kinase alpha
IKK: Inhibitor of nuclear factor Kappa-B Kinase complex
IL: InterLeukin
IRAK4: Interleukin Receptor-Associated Kinase 4
JAK: Janus kinase
KLHL: Kelch-like protein
LDLR: Low-Density Lipoprotein Receptor
LITAF: LPS-Induced TNFAlpha Factor
LPS: LipoPolySaccharide
MAPKs: Mitogen-Activated Protein Kinases
MBL: Mannose Binding Lectin
MGD1/2: Mytilus galloprovincialis Defensin 1 /2
MIF: Migration Inhibitory Factor
MNK: MAP kinase-interacting serine/threonine-protein kinase
MR1: Mannose Receptor 1

MyD88: Myeloid Differentiation primary response gene 88
NALPs: NATCH, LRR, and PYR containing proteins
NFkB: Nuclear Factor of kappa light polypeptide gene enhancer in B-cells
NLR: NOD-Like Receptor
NOD: Nucleotide Binding Oligomerization Domain
P13K: Phosphatidylinositol-4,5-bisphosphate 3-Kinase
PAC2: Proteasome Assembly Chaperone 2
PAMPs: Pathogen Associated Molecular Patterns
PGRP: Peptidoglycan Recognition Protein
PI31: Proteasome Inhibitor PI31 subunit
Pim: proto-oncogene serine/threonine-protein kinase Pim
PRR: Pathogen Recognition Receptors
RAB: Ras-related gtp-Binding protein
RIP: Receptor-Interacting serine-threonine kinase
ROS: Reactive Oxygen Species
SEC22: vesicle transport protein SEC22
SOD: SuperOxide Dismutase
STAT: Signal Transducer and Activator of Transcription protein
SRCR: Scavenger Receptor Cysteine-Rich protein precursor
TAK: mitogen activated protein kinase kinase
TIMP3: Tissue Inhibitors of MetalloProteinase 3
TNF: Tumour Necrosis Factor
TNFR: Tumour Necrosis Factor Receptor
TRAF6: TNF receptor-associated factor 6
Ub: Ubiquitin
UBR5: Ubiquitin protein Ligase E3 (component n-recognin 5)
α2: proteasome subunit alpha type 2
β4, β5: Proteasome subunit beta type 4/5

experimental measures. All the steps of the DNA microarray testing could be used to strengthen the final data interpretation, from the microarray preparation strategy to the stringency of the hybridization reaction to the algorithms applied to data processing.

The maintenance of the physical collection of the cDNAs, *i.e.*, recombinant bacterial clones, is a prerequisite for the use of spotted cDNA microarrays (for instance, the current use of Mytarray 1.0 slides, printed at the CRIBI facility depends on long work performed at the Department of Biology, University of Padua). Such work is not more affordable as long as the clustered ESTs increase in number, and external commercial services or deep sequencing become an attractive alternative.

As a matter of fact, next-generation sequencing (NGS) technologies are now complementing and challenging the DNA microarrays as alternative tools for genome analysis and transcriptome sequencing (Hurd and Nelson, 2009; Morozova *et al.*, 2009). For instance, the so called 454 pyrosequencing has been already applied to the study of tissue-specific expression patterns in *M. galloprovincialis* (Craft *et al.*, 2010) and many laboratories in the world are now investing in this kind of work.

References

Balseiro P, Falcó A, Romero A, Dios S, Martínez-López A, Figueras A, *et al. Mytilus galloprovincialis* Myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS One. 2011;6(8):e23140.

Banni M, Negri A, Mignone F, Boussetta H, Viarengo A, Dondero F. Gene Expression Rhythms in the Mussel *Mytilus galloprovincialis* (Lam.) across an Annual Cycle. PLoS One 6(5): e18904, 2011.

Ben Said O, Goñi-Urriza M, El Bour M, Aissa P, Duran R. Bacterial community structure of sediments of the Bizerte lagoon (Tunisia), a southern Mediterranean coastal anthropized lagoon. Microb. Ecol. 59: 445-456, 2010.

Carland TM, Gerwick L. The C1q domain containing proteins: Where do they come from and what do they do? Dev. Comp. Immunol. 34: 785-790, 2010.

Chen JJ, Hsueh HM, Delongchamp RR, Lin CJ, Tsai CA. Reproducibility of microarray data: a further analysis of microarray quality control (MAQC) data. BMC Bioinformatics 8: 412, 2007.

Costa F, Alba R, Schouten H, Soglio V, Gianfranceschi L, Serra S, *et al.* Use of homologous and heterologous gene expression profiling tools to characterize transcription dynamics during apple fruit maturation and ripening. BMC Plant Biol. 10: 229, 2010.

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.

Craft JA, Gilbert JA, Temperton B, Dempsey KE, Ashelford K, Tiwari B, *et al.* Pyrosequencing of *Mytilus galloprovincialis* cDNAs: Tissue-Specific Expression Patterns. Plos One 5(1): e8875, 2010.

Dodd RB, Drickamer K. Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. Glycobiology. 11(5):71R-79R, 2001.Dodds AW, Matsushita M: The phylogeny of the complement system and the origins of the classical pathway. Immunobiology 212: 233-243, 2007.

Evans TG, Somero GN. Phosphorylation events catalyzed by major cell signaling proteins differ in response to thermal and osmotic stress among native (*Mytilus californianus* and *Mytilus trossulus*) and invasive (*Mytilus galloprovincialis*) species of mussels. Physiol. Biochem. Zool. 83: 984-996, 2010.

GEO. The Gene Expression Omnibus can be accessed at http://www.ncbi.nlm.nih.gov/geo.

Gerdol M, Manfrin C, De Moro G, Figueras A, Novoa B, Venier P, *et al.* The C1q domain containing proteins of the Mediterranean mussel *Mytilus galloprovincialis*: a widespread and diverse family of immune-related molecules. Dev. Comp. Immunol. 35: 635-643, 2011.

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

Gibson G, Muse SV. A Primer of Genome Science. Sinauer Associates Inc., Sunderland, MA USA, 2004.

Gibson G. The environmental contribution to gene expression profiles. Nat. Rev. Genet. 9: 575-581 2008.

Gorbushin AM, Iakovleva NV. A new gene family of single fibrinogen domain lectins in *Mytilus*. Fish Shellfish Immunol. 30: 434-438, 2011.

Gracey AY, Chaney ML, Boomhower JP, Tyburczy WR, Connor K, Somero GN. Rhythms of gene expression in a fluctuating intertidal environment. Curr. Biol. 18: 1501-1507, 2008.

Hook SE. Promise and progress in environmental genomics: a status report on the applications of gene expression-based microarray studies in ecologically relevant fish species. J. Fish Biol. 77: 1999-2022, 2010.

Hurd PJ, Nelson CJ. Advantages of next-generation sequencing versus the microarray in epigenetic research. Brief. Funct. Genomic. Proteomic. 8: 174-183, 2009.

IMAQUANIM. Facts on this FP6 Integrated Project are at http://imaquanim.dfvf.dk/info.

Kerr MK, Churchill GA. Experimental design for gene expression microarrays. Biostatistics 2: 183-201, 2001.

Larsen PF, Schulte PM, Nielsen EE. Gene expression analysis for the identification of selection and local adaptation in fishes. J. Fish Biol. 78: 1-22, 2011.

Lenoir T, Giannella E. The emergence and diffusion of DNA microarray technology. J. Biomed. Discov. Collab. 1: 11, 2006.

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

Li H, Venier P, Prado-Alvárez M, Gestal C, Toubiana M, Quartesan R, et al. Expression of Mytilus immune genes in response to experimental challenges varied according to the site of collection. Fish Shellfish Immunol. 28: 640-648, 2010.

Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, et al. Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat. Biotechnology 14: 1675-1680, 1996.

Lockwood BL, Sanders JG, Somero GN. Transcriptomic responses to heat stress in invasive and native blue mussels (genus Mytilus): molecular correlates of invasive success. J. Exp. Biol. 213(Pt 20): 3548-3558, 2010.

Lockwood BL, Somero GN. Transcriptomic responses to salinity stress in invasive and native blue mussels (genus Mytilus). Mol. Ecol. 20: 517-29, 2011.

Malagoli D. Cytokine network in invertebrates: the very next phase of comparative immunology. Inv. Sur. J. 7: 146-148, 2010.

Manfrin C, Dreos R, Battistella S, Beran A, Gerdol M, Varotto L, et al. Mediterranean mussel gene expression profile induced by okadaic acid exposure. Environ. Sci. Technol. 44: 8276-8283, 2010.

Martín-Subero JI, Esteller M. Profiling epigenetic alterations in disease. Adv. Exp. Med. Biol. 711: 162-177, 2011.

MESH. Current MEdical Subject Headings can be accessed at http://www.ncbi.nlm.nih.gov/mesh.

Morozova O, Hirst M, Marra MA. Applications of new sequencing technologies for transcriptome analysis. Annu. Rev. Genomics Hum. Genet. 10: 135-151, 2009.

Nazar RN, Chen P, Dean D, Robb J. DNA chip analysis in diverse organisms with unsequenced genomes. Mol. Biotechnol. 44: 8-13, 2010.

NCT. The National Center for Toxicogenomics is at the National Institute of Environmental Health Sciences (http://www.niehs.nih.gov/research/atniehs/nct.cfm).

Padhi A, Verghese B. Molecular diversity and evolution of myticin-C antimicrobial peptide variants in the Mediterranean mussel, Mytilus galloprovincialis. Peptides 229: 1094-10101, 2008.

Pallavicini A, Costa MM, Gestal C, Dreos R, Figueras A, et al. High sequence variability of myticin transcripts in hemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev. Comp. Immunol. 32: 213-226, 2008.

Pantzartzi C, Drosopoulou E, Yiangou M, Drozdov I, Tsoka S, Ouzounis CA, et al. Promoter complexity and tissue-specific expression of stress response components in Mytilus galloprovincialis, a sessile marine invertebrate species. PLoS Comput. Biol. 6: e1000847, 2010.

Ptitsyn A, Schlater A, Kanatous S. Transformation of metabolism with age and lifestyle in Antarctic seals: a case study of systems biology approach to cross-species microarray experiment. BMC Syst. Biol. 4: 133, 2010.

Rogers S, Cambrosio A. Making a new technology work: the standardization and regulation of microarrays. Yale J. Biol. Med. 80: 165-178, 2007.

Romero A, Dios S, Poisa-Beiro L, Costa MM, Posada D, Figueras A, et al. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (Mytilus galloprovincialis) suggest ancient and complex immune recognition models in invertebrates. Dev. Comp. Immunol. 35: 334-344, 2011.

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 6(2):e17003, 2011.

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

Shaw JP, Dondero F, Moore MN, Negri A, Dagnino A, Readman JW, et al. Integration of biochemical, histochemical and toxicogenomic indices for the assessment of health status of mussels from the Tamar Estuary, UK Mar. Environ. Res. [in press] 2011.

Shi L, Perkins RG, Fang H, Tong W. Reproducible and reliable microarray results through quality control: good laboratory proficiency and appropriate data analysis practices are essential. Curr. Opin. Biotechnol. 19: 10-18, 2008.

Somero GN. Comparative physiology: a "crystal ball" for predicting consequences of global change. Am. J. Physiol. Regul. Integr. Comp. Physiol. 301: R1-R14, 2011

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. (NY) [in press] 2011.

Steinberg CE, Stürzenbaum SR, Menzel R. Genes and environment - striking the fine balance between sophisticated biomonitoring and true functional environmental genomics. Sci. Total Environ. 400: 142-161, 2008.

Thomas M, Lieberman J, Lal A. Desperately seeking microRNA targets. Nat. Struct. Mol. Biol. 17: 1169-1174, 2010.

Venier P, De Pittà C, Bernante F, Varotto L, De Nardi B, Bovo G, et al. MytiBase: a knowledge base of mussel (M. galloprovincialis) transcribed sequences. BMC Genomics 10: 72, 2009.

Venier P, De Pittà C, Pallavicini A, Marsano F, Varotto L, Romualdi C, *et al.* Development of mussel mRNA profiling: Can gene expression trends reveal coastal water pollution? Mutat. Res. 602: 121-134, 2006.

Venier P, Pallavicini A, De Nardi B, Lanfranchi G. Towards a catalogue of genes transcribed in multiple tissues of *Mytilus galloprovincialis*. Gene 314: 29-40, 2003.

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 12: 69, 2011.

Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC, *et al.* Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316: 1738-1743, 2007.

Yeaman MR, Yount NY. Unifying themes in host defence effector polypeptides. Nat. Rev. Microbiol. 5: 727-740, 2007.

Zhang SM, Zeng Y, Loker ES. Expression profiling and binding properties of fibrinogen-related proteins (FREPs), plasma proteins from the schistosome snail host *Biomphalaria glabrata*. Innate Immun. 14: 175-189, 2008.