



HAL
open science

The immune system of the freshwater zebra mussel, *Dreissena polymorpha*, decrypted by proteogenomics of hemocytes and plasma compartments

Maxime Lepretre, Christine Almunia, J. Armengaud, Arnaud Salvador, Alain Geffard, Mélissa Palos-Ladeiro

► To cite this version:

Maxime Lepretre, Christine Almunia, J. Armengaud, Arnaud Salvador, Alain Geffard, et al.. The immune system of the freshwater zebra mussel, *Dreissena polymorpha*, decrypted by proteogenomics of hemocytes and plasma compartments. *Journal of Proteomics*, 2019, 202, pp.103366. 10.1016/j.jprot.2019.04.016 . hal-02108964

HAL Id: hal-02108964

<https://hal.science/hal-02108964>

Submitted on 22 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The immune system of the freshwater zebra mussel, *Dreissena***
2 ***polymorpha*, decrypted by proteogenomics of hemocytes and**
3 **plasma compartments.**

4 Maxime Leprêtre^{1,3}, Christine Almunia², Jean Armengaud², Arnaud Salvador³, Alain Geffard¹,
5 Mélissa Palos-Ladeiro¹

6 ¹Université de Reims Champagne-Ardenne UMR-I 02 INERIS-URCA-ULH SEBIO Unité Stress
7 Environnementaux et, BIOSurveillance des milieux aquatiques, UFR Sciences Exactes et Naturelles,
8 Campus du Moulin de la Housse, BP 1039 51687 Reims CEDEX, France.

9 ²Laboratoire Innovations Technologiques pour la Détection et le Diagnostic (Li2D), Service de
10 Pharmacologie et Immunoanalyse (SPI), CEA, INRA, F-30207 Bagnols-sur-Cèze, France

11 ³Université de Lyon, Université Claude Bernard Lyon 1, Institut des Sciences Analytiques, CNRS
12 UMR 5280, F-69100 Villeurbanne, France.

13

14 Keywords: hemolymph, bivalve, immunity, non-model organism, proteogenomics

15

16 **Highlights**

- 17 • Proteogenomic analysis of the zebra mussel hemolymph
18 • Large-scale description of the molecular mechanisms involved in mussel immunity
19 • Discovery of immune proteins rarely observed in the hemolymph of bivalves

20

21 **Abstract**

22 The immune system of bivalves is of great interest it reflects the health status of these
23 organisms during stressful conditions. While immune molecular responses are well
24 documented for marine bivalves, few information is available for continental bivalves such
25 as the zebra mussel, *Dreissena polymorpha*. A proteogenomic approach was conducted on
26 both hemocytes and plasma to identified immune proteins of this non-model species.
27 Combining transcriptomic sequences with mass spectrometry data acquired on proteins is a
28 relevant strategy since 3,020 proteins were identified, representing the largest protein
29 inventory for this sentinel organism. Functional annotation and gene ontology (GO) analysis
30 performed on the identified proteins described the main molecular players of hemocytes
31 and plasma in immunity. GO analysis highlights the complementary immune functions of
32 these two compartments in the management of micro-organisms. Functional annotation
33 revealed new mechanisms in the immune defence of the zebra mussel. Proteins rarely
34 observed in the hemolymph of bivalves were pinpointed such as natterin-like and
35 thaumatin-like proteins. Furthermore, the high abundance of complement-related proteins
36 observed in plasma suggested a strong implication of the complement system in the immune
37 defence of *D. polymorpha*. This work brings a better understanding of the molecular
38 mechanisms involved in zebra mussel immunity.

39 Significance

40 Although the molecular mechanisms of marine bivalves are widely investigated, little
41 information is known for continental bivalves. Moreover, few proteomic studies described
42 the complementarity of both hemolymphatic compartments (cellular and plasmatic) in the
43 immune defence of invertebrates. The recent proteogenomics concept made it possible to
44 discover proteins in non-model organisms. Here, we propose a proteogenomic strategy with
45 the zebra mussel, a key sentinel species for biomonitoring of freshwater, to identify and
46 describe the molecular actors involved in the immune system in both hemocytes and plasma
47 compartments. More widely, this study provided new insight into bivalve immunity.

48

49 1. Introduction

50

51 Immune system acts as the main defence mechanism against invasive organisms. In
52 bivalves, the immune defences are mainly undertaken in the hemolymph compartment. This
53 circulatory fluid comprises immune cells, the hemocytes, and a cell-free compartment, the
54 plasma. Innate immune system process of bivalves can be summarized in three main steps:
55 (i) the recognition of molecular motifs associated with microorganisms (MAMPs) or
56 endogenous molecules secreted by damaged tissues (DAMPs) by soluble compounds and
57 cellular receptors, (ii) the activation of different signalling pathways, (iii) the production of
58 molecular effectors involved in host defense and cellular defence responses [1].

59 Due to their economic interest and the disease outbreaks faced in aquaculture,
60 molecular immune system of marine bivalves are well documented compared to continental
61 bivalves which are mainly used to assess water quality. Among continental bivalves, the
62 freshwater mussel, *Dreissena polymorpha*, represent the counterpart of the marine mussel
63 in ecotoxicological studies [2]. While cellular immune responses of hemocytes are well
64 characterized for *D. polymorpha*, few information are available on molecular immune
65 responses [3]. As a non-model species, no complete genome is currently available for *D.*
66 *polymorpha* and genetic markers are relatively scarce, especially for immune related genes.
67 To our knowledge, only a single study reported about thirty genes involved in hemocyte
68 responses against microbial antigens [4]. These genes, such as ATP synthase or Heat Shock
69 Protein 70, were more related to general stress responses than immune system defence.

70 To date, research on bivalve immunology mainly focused on immune responses of
71 hemocyte using transcriptomic investigations [5]. Because plasmatic compartment is
72 essentially mRNA-free, such approach does not give a comprehensive picture of the entire
73 hemolymph compartment. Regarding the proteome of the hemolymph of aquatic
74 invertebrates, the information remains scarce because only 2D gel electrophoresis results
75 have been till now reported [5]. This approach focuses only on the most abundant and
76 soluble proteins [6], explaining the relatively low number and low relevant proteins
77 identified in bivalve plasma.

78 Next-generation sequencing technologies have made great progress over the last
79 decade, providing information on i) structure and molecular function of the genetic material,
80 and ii) level of expression of genes. Proteins and metabolites are important components of
81 biological systems that can also be assessed at large scale by state-of-the-art tandem mass
82 spectrometry. The recent proteogenomics concept emerged as a relevant strategy for the
83 discovery of proteins in non-model organisms [7]. For animals and plants, this approach
84 combines data obtained by high-throughput mRNA sequencing (RNA-seq) for establishing
85 the most appropriate protein sequence database and large-scale protein shotgun
86 sequencing by tandem mass spectrometry. Indeed, peptide sequence identification relies on
87 the comparison of experimental peptide mass spectra obtained by tandem mass
88 spectrometry against all the possible theoretical spectra derived from the protein sequence
89 database. For non-model organisms whose genome have not yet been sequenced, it has
90 been proposed the construction of a theoretical protein sequence database on the basis of
91 the sequenced transcriptome *in silico* translated to amino acid sequences in all three
92 possible reading frames. Proteogenomics has already proven its capacity in quickly
93 documenting proteins of non-model sentinel species used for continental water
94 biomonitoring. For example, 1,873 proteins related to the reproductive system of *Gammarus*
95 *fossarum* were experimentally validated using proteogenomics [8].

96 In order to get insights into the immune proteome of the zebra mussel, proteogenomics
97 was conducted on both hemocytes and plasma compartment. For this, RNAseq and shotgun
98 mass spectrometry with a tandem mass spectrometer equipped with an ultra-high-field
99 Orbitrap analyzer were performed on the *D. polymorpha* sentinel species. Both data were
100 combined leading to the identification of 3,227 novel proteins. Proteins involved in immune
101 processes and potential immune proteins of both compartments were highlighted and
102 commented. This study contributes to a better understanding of the molecular mechanisms
103 involved in immunity of bivalves.

104

105 2. Material and Methods

106

107 2.1 Construction of a paired-end RNA-Seq library and illumina sequencing.

108

109 Zebra mussels, *Dreissena polymorpha*, between 15- and 25-mm long were collected at
110 the lake of Der (Grand Est, France) in July and September 2016. Dissection of i) gonad male
111 and ii) female, iii) gills, iv) digestive gland, v) whole tissues and sampling of vi) total
112 hemolymph was carried out directly after harvest. To avoid natural individual variation, six
113 organs of each type were pooled, frozen immediately in liquid nitrogen and stored at -80 °C
114 until total RNA extraction. Pools of hemolymph were centrifuged twice at 3,000 x g for 5 min
115 at 4 °C to separate the hemocytes from the plasma. Pellets of hemocytes were mixed with
116 500 µL of Trizol reagent (Sigma-Aldrich, Saint-Quentin Fallavier, France) in order to disrupt

117 cell walls. Tissue samples were disrupted with 1 ml of Trizol reagent (Sigma-Aldrich) per 100
118 mg of tissue and subjected to an ultra-turrax shredder run at 20,000 rpm for 5 min.

119 Total RNA of each sample was isolated according to Sigma-Aldrich instructions. Briefly,
120 samples were incubated at room temperature for 5 min. Then, a volume of 200 μ l of
121 chloroform was added to homogenize by shaking during 30 sec. After 15 min at room
122 temperature, samples were centrifuged at 12,000 x *g* for 15 min at 4 °C. The aqueous phase
123 was introduced into a new tube and 500 μ l of isopropanol was added. After 10 min of
124 incubation, samples were centrifuged at 12,000 x *g* for 8 min at 4 °C. RNAs were washed
125 with 1ml of 70% ethanol, followed by centrifugation at 7,500 x *g* for 5 min at 4 °C. RNAs were
126 dried, and then samples were resuspended in 40 μ l of RNase-free water and incubated in a
127 water bath at 57 °C for 15 min. Total RNAs for each sample were quantified using a Genova
128 Nano spectrophotometer (Jenway, UK) and their qualities were checked using agarose gel.
129 Samples were stored at -80°C until RNAseq library construction.

130 RNA Seq library construction and sequencing were performed by Acobiom (Montpellier,
131 France), using the HiSeq2000 sequencing instrument (Illumina). A total of six RNAseq
132 libraries, corresponding to each tissue sampled, were constructed according to the TrueSeq
133 Stranded (Illumina) protocol with 500 ng of total RNAs. Samples were sequenced using
134 paired-end reads 2x150 bp method and the reads were assembled with the software Trinity
135 pipeline (2.3.0 version) following the procedure of Haas et al. [9]. A transcriptome reference,
136 called "Dreissena_reference", was constructed by compiling the 6 RNAseq data.

137

138 2.2 Sample preparation for mass spectrometry analysis

139

140 Zebra mussels between 15- and 25-mm long were collected at the lake of Der (Grand Est,
141 France) in May 2017. Mussels were maintained for at least two weeks at the laboratory in 20
142 L tanks filled with spring water Cristaline Aurèle (Jandun, France), with a constant
143 temperature (14°C) and controlled aeration. Mussels were fed twice a week with a mixture
144 of two species of microalgae, *Chlorella pyrenoidosa* and *Scenedesmus obliquus* (two millions
145 of microalgae per day and per mussel). Tank water was removed weekly and replaced by
146 fresh water. After the acclimation period, 200 μ L of hemolymph per mussel was collected
147 from the posterior adductor muscle by gentle aspiration with a 30 G x 0.3 mL syringe. After
148 checking hemolymph quality under a binocular microscope, hemolymph samples were
149 pooled to reach 1 mL hemolymph per replicate, which represents a minimum of five and a
150 maximum of eight mussels per pool of hemolymph. Five biological replicates were
151 performed. Pools of hemolymph were then centrifuged at 800 x *g* for 15 min at 4 °C to
152 gently separate the hemocytes from the plasma. Hemocyte pellets were mixed with 500 μ L
153 of Tri-reagent (Sigma-Aldrich) and stored at -20°C until the protein extraction step while the
154 resulting supernatant (plasma) was immediately added to 3 volumes of ice-cold acetone and
155 stored overnight at -20 °C to allow protein precipitation. Precipitated proteins of plasma
156 were centrifuged at 15,000 x *g* for 10 min at 4 °C. The protein pellets were washed twice

157 with 1 mL of 80 % cold acetone followed by centrifugation at 15,000 x *g* for 10 min at 4 °C,
158 and removal of the supernatant. Then, the pellets were dried at room temperature and re-
159 suspended in 80 µL of re-suspension buffer consisting of 6 M urea; 2 M thiourea ; 4 % 3-[(3-
160 cholamidopropyl)dimethyl-ammonio]-1-propanesulfonate (CHAPS); 65 mM dithiothreitol
161 (DTT), 0.1 % Triton X100 (Vol:Vol), pH 7.8.

162 Hemocytes proteins were extracted following the procedure described by Vaibhav et al.
163 [10] with few modifications. Briefly, hemocytes were lysed by passing them ten times
164 through a syringe equipped with a 25 G needle and centrifuged at 12,000 x *g* for 2 min for an
165 initial clean up. RNAs were removed from the resulting supernatant by adding 100 µL of
166 chloroform, incubating for 15 min, centrifugation at 12,000 x *g* and 4 °C for 15 min, and
167 removing the colorless aqueous phase. DNA was discarded by adding 150 µL of absolute
168 ethanol for 3 min followed by centrifugation at 2,000 x *g* and 4 °C for 5 min. Then, proteins
169 from the resulting supernatant were precipitated by adding 3 volumes of ice-cold acetone to
170 the supernatant. Samples were incubated at room temperature for 10 min and then
171 centrifuged for 10 min at 12,000 x *g* and 4 °C. The remaining protein pellet was then washed
172 three times for 10 min in 1 mL of 0.3 M guanidine hydrochloride in 95 % ethanol (Vol:Vol)
173 followed by centrifugation at 8,000 x *g* and 4 °C for 5 min and removal of the supernatant. A
174 final wash was performed with 1 mL of 95 % ethanol before protein pellets were dried at
175 room temperature and re-suspended in 80 µL of resuspension buffer.

176 The protein concentration of each sample was measured by Bradford assay with a
177 calibration curve established with bovine serum albumin (BSA). Biological replicates were
178 diluted in re-suspension buffer to obtain the same protein concentration of 2.25 µg/µL. A
179 volume of 20 µL of each sample was diluted with 10 µL of LDS 3X buffer (Invitrogen)
180 consisting of 80 mM Tris / HCl, 106 mM Tris base, 1.5 % lithium dodecyl sulfate, 7.5 %
181 glycerol, 0.38 mM EDTA, 0.17 mM SERVA Blue G250, 0.131 mM phenol red, buffered at pH
182 8.5 and supplemented with 2.5 % beta-mercaptoethanol. Then, samples were boiled for 5
183 min at 95 °C, centrifuged briefly, and loaded onto a 4–12 % gradient 10-well NuPAGE gel
184 (Invitrogen). A 5 min electrophoretic migration of the proteins was performed at 200 V in
185 MES buffer (Invitrogen). After Coomassie Blue Safe staining (Invitrogen) and destaining with
186 ultra-pure water, each polyacrylamide band corresponding to each whole proteome was
187 excised. Proteins from each polyacrylamide band was first alkylated with iodoacetamide,
188 reduced with DTT and then proteolyzed with Sequencing Grade Trypsin (Roche) in presence
189 of 0.01 % ProteaseMAX surfactant (Promega) for 1 h at 50 °C as previously described by
190 Hartmann et al. [11]. The proteolysis was then stopped with the addition of TFA 5 % to reach
191 a final concentration of 0.5 %.

192

193 2.3 NanoLC-MS/MS analysis

194

195 The resulting peptide mixtures were analyzed in a data-dependent mode with a Q
196 Exactive HF tandem mass spectrometer (ThermoFischer Scientific) coupled on line to an
197 Ultimate 3000 chromatography system (ThermoFisher), essentially as previously described
198 [12]. A volume of 5 μ L of each sample was injected, first desalted with a reverse-phase
199 capillary precolumn C₁₈ PepMap TM 100 (LC-Packing) and then resolved with a nanoscale
200 500 mm C18 PepMap TM 100 (LC-Packing) column using a 90 min gradient from 20 % to 80
201 % of CH₃CN, 0.1 % formic acid. Mass spectra of peptide ions with 2+ and 3+ charge states
202 were acquired in the full scan range m/z 350-1800, with a resolution set at 60,000, an AGC
203 target at 3×10^6 , following a Top20 method for selecting peptides for fragmentation. MS/MS
204 mass spectra were acquired with an AGC target set at 10^5 , a loop count of 60 msec, an
205 isolation window of 1.6 m/z , a resolution at 15,000, and a scan carried out in a mass range
206 from 200 to 2,000 m/z .

207

208 2.4 Proteogenomics database design and MS/MS interpretation

209

210 The Dreissena_reference transcriptomic database was translated into the six possible
211 reading frames with a homemade python script. The resulting translated contig database
212 contained 5,648,820 polypeptide sequences, totaling 941,472,234 residues. MS/MS spectra
213 were assigned to peptide sequences with the Mascot search engine with the following
214 parameters: full-trypsin specificity, a maximum of two missed cleavages, a mass tolerance of
215 5 ppm on the parent ion and 0.02 Da on the MS/MS analysis, and a pvalue for peptide
216 validation at 0.1. Modifications taken into account were carboxyamidomethylated cysteine
217 (+57.0215) as static modification and oxidized methionine (+15.9949) and deamidation of
218 asparagine and glutamine (+0.9848) as dynamic modifications. Translated contigs identified
219 with at least one MS/MS certified peptide were selected for creating a reduced database
220 named "Dreissene_cascade". This second database comprised 54097 polypeptide
221 sequences, totaling 20,419,358 residues. Then, a second Mascot round of search was
222 performed with a p value threshold for validating peptide assignment set at 0.05. At this
223 stage, a protein was validated when at least two unique peptide sequences were detected in
224 at least one of the ten samples used in this study. The false positive rate for protein
225 identification was estimated, through a search with a reverse decoy database, as less than
226 1% using the same parameters. For each protein and for each replicate, the numbers of
227 spectral counts observed for the different detected peptides and corresponding to a
228 peptide-to-spectrum assignation were summed. Then, the spectral counts detected in each
229 of the five biological replicates were merged to obtain the spectral counts of each protein in
230 each of the two conditions, hemocytes and plasma. As previously proposed [13], the
231 normalized spectral abundance factor (NSAF) of each protein was calculated as the sum of
232 the spectral counts (SC) observed in the five replicates of hemocytes or plasma, divided by
233 its molecular mass expressed in kDa. The NSAF values were then presented as a ratio
234 (%NSAF) by dividing the NSAF of each protein to the sum of all NSAF values detected in

235 hemocytes or plasma. Spectral count missing values are observed in the case of low
236 abundant proteins, no specific correction of missing values was made.

237

238 2.5 Protein functional annotation and gene ontology analysis

239

240 Functional annotation of MS/MS identified proteins was derived from a sequence
241 similarity search without taxonomical restriction against (i) the National Center for
242 Biotechnology Information (NCBI) database of non-redundant protein sequences (nr)
243 (131,025,392 sequences released 05-09-2017) and (ii) the Swissprot database (465,089
244 sequences released, 30-09-2016), using the blastp module of DIAMOND (version 0.8.22) with
245 an E-value threshold set at 1E-03. Functional annotations were further confirmed by
246 searching conserved domain (CD) of proteins with the Batch CD-Search tool of NCBI server
247 [14]. Gene Ontology (GO) assignment for each translated contig was obtained with the GO
248 database (<http://www.geneontology.org/>). GO assignments were used to classify molecular
249 processes and biological processes of the hemocyte and plasma proteins. A focus on
250 immune-related proteins was performed by studying GO annotation of the proteins that
251 matched with the GO-term “immune system process”.

252 2.6 Mass spectrometry and proteomics data

253

254 The mass spectrometry and proteomics data have been deposited to the
255 ProteomeXchange Consortium via the PRIDE [15] partner repository with the dataset
256 identifier PXD012991 and 10.6019/PXD012991.

257

258 3. Results and discussion

259

260 3.1 A double round search for proteogenomics identification of proteins from the 261 hemolymph of *D. polymorpha*.

262

263 We applied a proteogenomics-derived strategy for quickly identifying the proteins from
264 two fractions of the hemolymph of *Dreissena polymorpha* physiologically acclimated to
265 laboratory conditions. Figure 1 shows a schematic representation of the experimental
266 strategy and the double round searches for interpreting the tandem mass spectrometry
267 data. First, RNAseq was performed on six samples of mRNAs extracted from six different
268 tissues of zebra mussels: i) total hemolymph, ii) gonad male and iii) female, iv) gills, v)
269 digestive gland and vi) whole tissues, as indicated in Figure 1. A total of 160,869,134 paired-
270 end reads of 150 bp were assembled into transcripts with the Trinity software resulting in a
271 list of 941,470 contigs with a N50 of 594. A global proteogenomics reference database
272 “*Dreissena_reference*” was created consisting of the systematic translation of all these

273 contigs into the six possible frames. A dataset that will be described elsewhere (Lepretre et
274 al., unpublished data) and comprising 3,211,276 MS/MS spectra recorded on the proteomes
275 of various hemocytes and plasma samples was interpreted against the
276 “dreissena_reference” RNAseq-derived database to select relevant contigs for creating a
277 reduced database as previously recommended [16]. The attribution rate was 30% for
278 hemocytes and 29% for plasma, a result similar to other proteogenomics analysis performed
279 on non-model organisms with inflated RNA-seq proteogenomics database [17]. The resulting
280 reduced database comprised 54,097 protein sequences, a size comparable to classical
281 animal proteomics based on a sequenced and annotated genome. For the present study, five
282 biological samples of hemolymph were treated to get five fractions of hemocytes and five
283 fractions of plasma. Their protein content was explored by next-generation shotgun
284 proteomics resulting in a total of 281,799 MS/MS spectra recorded for the hemocytes and
285 261,188 MS/MS spectra for the plasma. As indicated in Figure 1, the second round of search
286 resulted in the interpretation of 103,979 MS/MS spectra (37%) and produced a list of 22,578
287 unique peptide sequences leading to the discovery of a total of 3,020 proteins in the
288 hemolymph (hemocytes and plasma), validated by at least two unique peptide sequences.

289 A higher number of proteins were identified in hemocytes compared to plasma (Fig. 2A).
290 1,489 proteins were assigned exclusively to the hemocytes and 840 proteins were observed
291 only in the plasma compartment while 691 proteins were identified in both compartments.
292 Separation process of hemocytes from plasma by centrifugation may produce cell lysates
293 leading to a protein contamination of plasma. In this case, some of the proteins identified
294 with a higher NSAF% in hemocytes compared to plasma could be exclusive to hemocytes.
295 433 proteins were following this trend, which represent 60% proteins of the 691 proteins
296 observed in both hemocytes and plasma. Because the same efforts in terms of tandem mass
297 spectrometry have been done for the two kinds of samples, results are probably directly
298 correlated to the dynamic range of these proteomes. A similar trend was observed in the
299 proteome of hemolymph from the marine mussel *Mytilus edulis* [18]. The authors identified
300 a total of 654 proteins in the hemolymph, 321 proteins were observed exclusively in
301 hemocytes, 178 were exclusively assigned to the plasma and 178 were common to both
302 compartments. Functional annotation of these proteins was performed by sequence
303 similarity search against the NCBI nr and Swissprot databases. Most of the proteins (2,948)
304 could be functionally annotated, but a relatively small set of protein sequences (2 %) did not
305 present any significant sequence similarities with known proteins and thus could be
306 considered as orphan proteins. Regarding the species origin of the most similar BLAST hit
307 against NCBI nr, the majority of proteins showed as expected strong sequence similarities
308 with proteins from molluscs taxa (Fig. 3). The pacific oyster *Crassostrea gigas* is the most
309 phylogenetic closely-related species based on this comparison with a total of 750 proteins
310 with high similarities to *D. polymorpha* MS/MS-detected proteins, followed by another
311 marine bivalve *Mizuhopecten yessoensis*. Due to their economic importance in aquaculture,
312 these two bivalves are highly studied and a large number of genomic resources are available,
313 making them the closest genetic resources to the zebra mussel known to date [19]. As

314 expected, *D. polymorpha* was poorly represented with only 8 proteins as its genome is not
315 yet sequenced and only a restricted set of proteins are till now present in the NCBI nr
316 database.

317

318 3.2 Gene ontology analysis of the immune proteome

319

320 A Gene ontology analysis was performed on the whole set of mass spectrometry certified
321 proteins in order to describe biological processes and molecular functions of the zebra
322 mussel's proteins from the hemolymph compartment. Focusing on the immune system, 563
323 proteins in hemocytes against 490 in plasma were assigned to the GO-term "Immune system
324 process" (Fig. 2B). Among them, 216 immune-related proteins were observed in both
325 hemocytes and plasma compartments, representing almost half of the immune proteins
326 identified in plasma. To understand the role of hemocyte and plasma compartments in *D.*
327 *polymorpha* in the management of immune responses, a gene ontology (GO) analysis was
328 performed on the set of proteins mapped with the GO-term "immune system process",
329 which were considered as immune proteins (Fig. 4). For this proteomics analysis, we choose
330 a label-free comparison of the two compartments based on spectral count as this method is
331 relatively independent of the ionizability variability of peptides. Furthermore, we
332 systematically normalized the spectral counts taking into account the mass of each
333 polypeptides, thus allowing a comparison of the abundances of various proteins. In this
334 context, Go-term annotations were weighted with NSAF values in order to compare both
335 compartments in terms of abundances.

336 Results of biological process Go-terms showed a higher NSAF value of immune proteins
337 of plasma (0.72) compared to hemocytes (0.45). While overall the same trend of biological
338 process and molecular function was observed between the hemocyte and plasma
339 compartments, few differences could be noticed (Fig. 4). The most represented biological
340 process of immune proteins was "response to stress", with a NSAF twice as important in
341 plasma (0.66) compared to hemocytes (0.32). Such results confirm that immune responses
342 occurring in plasma and hemocytes are highly sensitive to stressful conditions, making them
343 a relevant marker for ecotoxicological studies. Another highly represented biological process
344 was "vesicle-mediated transport" in both compartments. Exocytosis of immune proteins
345 from the hemocytes to extracellular region plays a critical role in the elimination of foreign
346 particles. This process occurs principally through vesicle-mediated transport, which mediates
347 the release of proteins in the extracellular space that participate in immune signalling,
348 recognition and destruction of micro-organisms [20]. The biological process "signal
349 transduction" was also prominent in both hemocyte and plasma compartments. Signal
350 transduction is required for the modulation of immune responses. This process is initiated by
351 the recognition of MAMPs or DAMPs by cellular receptors and/or soluble compounds
352 localized in cells and extracellular compartments. Immune signalling process plays a central

353 role by controlling effective immune responses against microorganisms. Immune signalling
354 pathways have been largely documented in hemocytes of bivalve species [21]. In plasma,
355 immune signal transduction also plays important role in controlling the defence responses of
356 bivalves. For example, in molluscs, the extracellular signal-regulated kinase (ERK) signalling
357 pathway regulates many immune process of hemocytes, including inflammation,
358 proliferation and apoptosis [22]. In both compartments, the GO analysis revealed a high
359 NSAF value of immune proteins related to membrane organization, cell motility and cell
360 adhesion. Recruitment and migration of hemocytes to the site of infection is one of the key
361 steps of the immune response. These processes involve a dynamic cytoskeleton remodeling
362 and shape changes of hemocytes, mediating cell motility, cell adhesion and
363 endocytosis/phagocytosis of foreign particles [23]. Nitrogen compound metabolic process
364 and cell death process were also highly represented in both compartments. This is not
365 surprising since molluscs produce nitrogen species for their basal metabolism as well as in
366 response to exogenous stimuli and programmed cell death is involved in homeostasis of the
367 immune system [24]. In hemocytes, a higher NSAF of immune proteins was observed for the
368 GO-terms related to cytoskeleton organization, cell cycle and autophagy, this latter is
369 considered as an important immune process of bivalve's hemocytes [25]. Interestingly, some
370 immune proteins were related to neurological system process in both compartments. While
371 the nervous system of invertebrates is poorly studied, Malagoli and Ottaviani [26] evidenced
372 a strong complementarity between the nervous system and immune system of mollusc
373 species.

374 Regarding molecular functions of immune proteins, more differences were observed
375 between the two compartments (Fig. 4B). Molecular function related to peptidase activity
376 was the most represented in the hemolymph with a NSAF value four times higher in plasma
377 compared to hemocytes. Similarly, immune proteins related to enzyme regulator activity
378 were more represented in plasma than hemocytes (0.34 in plasma and 0.08 in hemocytes).
379 Proteolytic enzymes are considered as integrative components of the bivalve defence
380 system [1]. They are involved in various immune system processes, such as immune
381 signalling, cell death and degradation of microbial cell wall components [27,28]. The plasma
382 compartment of bivalves is an important source of proteolytic enzymes that govern immune
383 responses. For example, proteolytic cascades are essential to the activation of immune
384 responses occurring in the plasma of bivalves, such as the complement system and the
385 prophenoloxidase systems [1]. Finally, immune proteins with functions associated with
386 general cell metabolism were more represented in hemocytes compared to plasma. In this
387 group are included proteins assigned to ATPase activity, phosphatase activity, and ligase
388 activity (Fig. 4B).

389

390 3.3 Identification of immune-related proteins of the hemolymph

391

392 According to the literature, several proteins were identified as potentially involved in the
393 immune defence of zebra mussel. In this section, a focus will be given on the proteins
394 involved in the (i) recognition of microorganisms, (ii) immune signalling, (iii) micro-organisms
395 internalization, (iv) elimination/destruction of microorganisms, and (v) apoptosis. It is
396 important to highlight that the proteins described in this chapter are proteins known to be
397 involved in the immune defence of organisms different than zebra mussels. Further
398 investigations with mussels challenged with immuno-modulators are necessary to validate
399 their involvement in *D. polymorpha* immunity. Whereas most of the immune proteins were
400 shared between hemocyte and plasma compartments some of them appeared to be specific
401 of one compartment. Informations on the functional annotation and abundance of identified
402 proteins in hemocytes and plasma are detailed in the supplement Table S1 and Table S2
403 respectively. Supplement Table S3 details conserved domain (CD) of proteins, which
404 supports the functional annotation.

405

406 3.3.1 Recognition of micro-organisms

407

408 Several proteins acting as pattern recognition receptors (PRRs) were found in high
409 abundance in the hemolymph of *D. polymorpha*. In bivalves' immunity, PRRs are often
410 associated with host defence responses against infection [1]. They allow the recognition of
411 MAMPs or DAMPs, which lead to the activation of different immune processes. Interestingly,
412 a natterin-like protein containing a jacalin-lectin domain and a pore-forming toxin domain
413 was identified as a major protein in the zebra mussel plasma (898 SCs) and hemocytes (804
414 SCs), which represents about 2.5% of the proteome in both compartments (Table 1). To our
415 knowledge, this is the first time that a natterin-like protein is identified in the hemolymph of
416 invertebrates. Natterins were first described as the major toxins from the venomous fish
417 *Thalassophryne nattereri*. According to the literature, this proteolytic enzyme could play an
418 important role in the innate immune process. Firstly, a kinogenase activity was characterized
419 in natterins purified from *T. nattereri* [29]. Further investigations showed that natterins exert
420 cytotoxic effects and can induce cell necrosis by degrading or remodelling extracellular
421 matrix [30]. Looking at its structure, a pore-forming aerolysin domain and a N-
422 terminal domain were identified in natterin genes isolated from the blood of lamprey
423 (*Lampetra japonica*) and the skin mucus of the Atlantic cod (*Gadus morhua*) [31,32]. Taken
424 together, these findings suggest that natterins act as pathogen recognition proteins with
425 potential toxin function.

426 In total, 4,962 SC were attributed to proteins belonging to the lectin family in the *D.*
427 *polymorpha* hemolymph (Table 1). Made of carbohydrate recognition domains (CRD), lectins
428 enable the recognition of the "non-self" carbohydrate patterns localized on the surface of
429 microorganisms such as glycans of parasite, bacteria lipopolysaccharide or glycoproteins of
430 the viral envelope, leading to the internalization and destruction of invaders by phagocytic
431 cells [33]. Lectins have been classified into several families depending on their structural fold

432 of carbohydrate recognition domain. In *D. Polymorpha*, C-type lectins (CTLs), R-type lectins,
433 galectins and fibrinogen-related proteins (FREPs) were found in both hemocytes and plasma
434 while a P-type lectin was exclusively found in plasma. CTLs were five times more abundant in
435 plasma (956 SCs) in comparison to hemocytes (169 SCs) (Table 1). The function of CTLs in the
436 immune system of molluscs has been firmly established. Several studies have demonstrated
437 the involvement of CTLs in the recognition, immobilization and opsonization of
438 microorganisms in the plasma and hemocytes of many mollusc species [34–36].
439 Furthermore, the expression of CTLs in hemocytes was modulated during an infectious
440 experiment, which confirms their involvement in the defence against pathogens [37].
441 Mytilecs proteins are R-type lectins isolated from the Mediterranean mussel, *Mytilus*
442 *galloprovincialis*. The structure of Mytilecs is composed of a glycan-binding domain and a
443 pore-forming aerolysin-like domain [38]. Like natterins, in addition to its pathogen
444 recognition function, the aerolysin-like domain of Mytilec are potentially directly devoted to
445 kill micro-organisms [39]. Two proteins of the zebra mussel hemolymph showed sequences
446 similarities with Mytilecs proteins. One was observed in high abundance in both hemocytes
447 and plasma while the other Mytilec, which contains an aerolysin-like domain, was only found
448 in plasma. Galectins were also observed in high abundance in the intra- and extra-cellular
449 compartments of zebra mussel hemolymph (588 SCs in hemocytes and 484 SCs in plasma)
450 (Table 1). Invertebrate galectins have been shown to play crucial roles in innate immunity.
451 In addition to their role in the recognition of MAMPs, galectins participate to many immune
452 processes such as cell adhesion, phagocytosis, immune regulation and pathogen destruction
453 [40–42]. 222 and 428 spectra associated with proteins containing fibrinogen-related
454 domains (FreDs) were identified in hemocytes and plasma respectively. While these lectin-
455 like molecules are known to have coagulation functions in vertebrates, the primary role of
456 invertebrates FreDs is immune recognition of microorganisms [43]. For example, FreDs in the
457 hemolymph of the bay scallop *Argopecten irradians* were significantly up-regulated after a
458 gram-negative bacteria infection and presented agglutinating properties, emphasizing the
459 contribution of FREPs to non-self recognition in invertebrates [44]. Other lectins type such as
460 mannose-6-phosphate-receptor were identified exclusively in plasma (57 SCs) (Table 1).
461 Cation-dependent mannose-6-phosphate receptor (MPR) are P-type lectin acting as PRR.
462 Zhang et al. [45] showed that MPR of shrimp was able to recognize invading bacteria and
463 stimulate the production of antimicrobial peptides in the hemolymph of the shrimp
464 *Marsupenaeus japonicus* during *in vivo* challenges with gram-positive and gram-negative
465 bacteria.

466 Other immune PRRs, were found in the hemolymph of *D. polymorpha*. A Beta-1,3-glucan-
467 binding protein (GBP) was identified in both compartments of the hemolymph of *D.*
468 *polymorpha* with more than 200 spectra counted in both compartments and a
469 peptidoglycan-recognition protein (PGRP) was found ten times more abundant in plasma
470 compared to hemocytes (Table 1). GBP act as a PRR by recognizing β -1, 3-glucan molecules
471 of the bacteria cell wall and trigger the innate immune responses [46]. PGRP are secreted
472 proteins which have been characterized in the hemolymph of oyster and scallops, and their

473 expression were found enhanced after bacteria challenges [47,48]. This PRR has a lysozyme
474 domain, which may confer bacteriolytic activity in addition to the peptidoglycan binding
475 capacity [47].

476

477 3.3.2 Proteins involved in immune signaling

478

479 Immune signalling is a crucial step in adapting an effective immune response leading to
480 the destruction of different kinds of microorganisms. Few proteins involved in immune
481 signalling of *D. polymorpha* were observed in plasma compared to hemocytes (Table 1). Of
482 them, two proinflammatory cytokines named allograft inflammatory factor 1 (AIF-1) and
483 macrophage migration inhibitory factor (MIF) were identified with a higher abundance in
484 hemocytes (84 SCs) compared to plasma (43 SCs) (Table 1). AIF-1 and MIF have been
485 identified as regulator molecules of the immune system of bivalve species. MIFs were
486 modulated upon bacterial challenge in the hemocytes of the clam *Ruditapes philippinarium*
487 and the scallop *Chlamys farreri* [49,50]. Both studies concluded that MIF cytokines were
488 involved in proinflammatory responses by promoting hemocyte migration to the wound
489 healing site. AIF-1 was also up-regulated under several microorganism challenges and tissue
490 injury in the hemocytes of bivalves [51–53]. Like in vertebrates, AIF-1 of bivalves promoted
491 inflammatory processes and stimulated phagocytosis of foreign microorganisms.

492 Mitogen-Activated Protein Kinases (MAPKs) are important proteins involved in the
493 regulation of cellular processes including cell proliferation, apoptosis and immune response,
494 especially against virus, fungi and bacteria [54]. Till now, three major MAPKs pathways have
495 been identified in the hemolymph of bivalves, the c-jun N-terminal kinases (JNKs), p38
496 mitogen-activated protein and the extracellular signal-regulated kinases (ERKs) [33]. Several
497 studies showed the implication of these MAPKs pathways in the regulation of immune signal
498 transduction, leading to the production of immune effectors adapted to the infectious
499 pathogen [22,55,56]. In the hemolymph of *D. polymorpha*, MAPKs were evidenced by 320
500 spectra in hemocytes and 80 spectra in plasma (Table 1). Key components of ERKs (ERK1,
501 ERK2) and p38 mitogen-activated protein (p38-MAPK and MAP2K-6) were identified in
502 hemocytes and plasma, while a component of the JNK pathway (stress-activated protein
503 kinase JNK) was exclusively identified in hemocytes.

504 The interferon signalling pathway is considered as a powerful antiviral defence of the
505 innate immune system [57]. While the main anti-viral defence in bivalves was associated
506 with the blocking of viral replication by RNA interference, recent studies evidenced
507 interferon (IFN) signalling pathway similar to vertebrates to counteract virus invasion [58–
508 60]. Briefly, the recognition of double-strand virus by PRRs leads to the activation and
509 expression of interferons, resulting to the activation of signal transducers and activators of
510 transcription (STATs), which stimulate the production of antiviral effectors such as the
511 interferon-induced protein 44-like (IFN-44) [58,61]. In our study, IFN-44 and a STAT-related

512 protein were exclusively identified in the hemocytes of the zebra mussel with 97 SCs. Their
513 detection suggests the involvement of IFN pathway in antiviral defences of *D. polymorpha*
514 (Table 1).

515 Other proteins involved in immune signalling, such as the nuclear factor kappa B (NF- κ B),
516 protein kinase C -type (PKC) and phosphatidylinositol 4,5-bisphosphate 3-kinase (PI3K), were
517 identified in the hemolymph of *D. polymorpha*. The transcription factor NF- κ B was already
518 reported to be present in hemocytes of many bivalve species and was considered as a key
519 regulator of the immune response by its participation in the production of immune effectors,
520 immune cytokines and proinflammatory factors [33]. The protein kinase PKC and PI3K are
521 important regulators of phagocytosis process of bivalve hemocytes [62].

522

523 3.3.3 Proteins involved in the internalization of microorganisms

524

525 Internalization of microorganisms by hemocytes is a fundamental process of innate
526 immune defence, causing the destruction of invaders. Internalization process consists of
527 three main stages: (i) the cell migration towards the microorganisms, (ii) cell adhesion and
528 (iii) internalization. While these three main stages are well characterized at the cellular level
529 in many invertebrates, little is known on the molecular actors that govern these responses.
530 To date, most of the proteins known to be involved in internalization processes are related
531 to the cytoskeletal regulatory proteins [63]. Indeed, cell migration, cell adhesion and
532 internalization require actin polymerization, which generates mechanical forces and
533 morphological changes to meet and ingest invaders [64].

534 Several studies have highlighted the central role of the actin-related protein 2/3 complex
535 (Arp2/3) in actin filament assembly during migration and phagocytosis process of
536 macrophages [64–66]. The activation of the Arp2/3 complex is carried out locally under the
537 action of proteins of the WASP family (Wiskott-Aldrich syndrome protein). These latter are
538 activated at the cell membrane by small GTPases of the Rho family such as Cell Division Cycle
539 42 (CDC42) and Ras-related C3 botulinum toxin substrate (Rac1). In invertebrates, the
540 implication of Arp2/3 complex in cell motility and phagocytosis by hemocytes have been
541 evidenced in *Drosophila* and oyster [65,69]. For example, Lau et al. [67] observed an up-
542 regulation of proteins involved in the regulation of arp2/3 complex associated with an
543 increase in hemocyte motility when oysters were challenged with the parasite *Perkinsus*
544 *marinus* and latex beads. In the hemolymph of *D. polymorpha*, many proteins involved in the
545 regulation of the Arp2/3 complex were identified. CDC42, RAC1, and Arp2/3 were detected
546 in both compartments with higher abundance in hemocytes compared to plasma while
547 neural Wiskott-Aldrich syndrome protein (NWASP) was exclusively observed in the
548 hemocytes.

549 Integrin subunits alpha and beta were eleven times more represented in hemocytes (941
550 SCs) compared to plasma (85 SCs) (Table 1). Probably, their presence in plasma may be the

551 result of plasma contamination by hemocytes during centrifugation process. These
552 transmembrane adhesion receptors play critical roles in the innate immune process such as
553 cell migration, cell attachment with pathogens, phagocytosis and apoptosis of immune cells
554 (reviewed in [68]). The implication of integrins in the internalization of microorganisms has
555 been documented in several invertebrate species, including bivalves [69–71]. For example,
556 integrins were up-regulated during bacterial infection in the Pacific oyster hemocytes [72].
557 The authors showed that the phagocytosis rate of bacteria decreased significantly after
558 integrin blocking with antibodies, which evidences the involvement of integrins in the
559 internalization of pathogens.

560 Proteins with functions related to clathrin-mediated endocytosis, including clathrin
561 heavy chain and adaptor proteins (APs), were produced mainly in the zebra mussel
562 hemocytes (349 SCs) (Table 1). Cellular uptake by the clathrin-mediated endocytosis is
563 crucial for many cellular functions such as nutriment uptake, cell homeostasis and growth
564 control [73]. Moreover, the clathrin-mediated pathway is used by viruses to enter into the
565 cells [74,75]. Binding of viral particles to cell receptors triggers the accumulation of key
566 components of clathrin-mediated endocytosis including clathrin, AP2 and dynamin, which
567 leads to clathrin coat assembly and cargo internalization [75]. The clathrin-mediated
568 endocytosis of viruses has also been demonstrated in the hemocytes of *Marsupenaeus*
569 *japonicus* and *Penaeus monodon* shrimps [76,77].

570 Other proteins with functions related to cytoskeleton remodeling and cell motility were
571 detected in the zebra mussel hemocytes. In this group we included dedicator of cytokinesis
572 protein 1, myosin II and Rho-associated protein kinase 2.

573

574 3.3.4 Humoral Immune Effectors

575

576 Humoral Immune effectors, antimicrobial peptides, proteases and protease inhibitors are
577 viewed as key elements of the innate immunity of bivalves since they participate to the
578 degradation of specific substrates including microbial cell walls [1].

579 As the first line of defence of the innate immune system of shellfish, antimicrobial
580 peptides (AMPs) are immediate, rapid and powerful weapons to protect organisms against
581 invasive pathogens. A thaumatin-like protein (TLP), containing a thaumatin domain, was the
582 most abundant protein considered as AMP in both hemocytes (201 SCs) and plasma (149
583 SCs) of *D. polymorpha* (Table 1). This antifungal protein has been discovered in a wide range
584 of organisms, including plants, fungi and insects [78]. TLPs are believed to exert antifungal
585 activity using a variety of enzymatic activities, which make them a potent AMP. Indeed, TLPs
586 have been shown to exert glucan binding and β -1,3-glucanase activities in addition to their
587 function as glucose polymers degrading, chitinase, xylanase, as well as α -amylase
588 inhibiting activities [78]. For bivalves, information concerning TLPs is still very limited. To our
589 knowledge, TLPs were only identified in hemocytes of the freshwater Pearl mussel (*Hyriopsis*

590 *cumingii*) and in mucosal secretions in the eastern oyster, *Crassostrea virginica* [79,80]. A
591 Defensin-like protein was identified exclusively in the hemocytes of the zebra mussel (Table
592 1). Defensins are a family of antimicrobial peptides found in many organisms, including
593 plants, vertebrates and invertebrates [81]. These are small cationic peptides which are
594 widely involved in innate immunity due to their antimicrobial activity [81]. Defensins have
595 been identified in several bivalve species and were shown to be active against a wide range
596 of bacteria, protozoan and fungi [82,83]. A defensin-like protein had already been identified
597 in tissues of the zebra mussel, including hemocytes [84]. Authors showed that the expression
598 of the *D. polymorpha* defensin from hemocytes exerted antimicrobial activity against gram-
599 negative bacteria during *in vitro* challenges. Other proteins considered as part of the
600 antimicrobial arsenal such as lysozymes, bactericidal/permeability-increasing protein (BPI)
601 and histones were also observed in the hemolymph of *D. polymorpha*. Lysozymes were
602 mainly identified in plasma with 51 SCs (Table 1). Two type of lysozymes were observed, a C-
603 type lysozyme (Chicken type) lysozyme and an I-type lysozyme (Invertebrates-type)
604 containing a destabilase domain. Lysozymes are part of the family of glucoside hydrolases,
605 they are found in a wide variety of tissues and secretion from vertebrates to invertebrates
606 [85]. The bacteriolytic enzymes I-type lysozymes are well documented in invertebrates and
607 have been found to play a major role in the destruction of bacteria due to their antimicrobial
608 activities [85]. In contrast, the functions of C-type lysozymes have been poorly studied in
609 bivalves [85]. To date, C-type lysozymes were only identified in tissues of the mussel and
610 displayed lytic activities against Gram-positive and Gram-negative bacteria [86]. BPI protein
611 was exclusively observed in the hemocytes of the zebra mussel with a low abundance. BPIs
612 are strong pore-forming agents exerting antibiotic action towards Gram-negative bacteria. A
613 BPI has been characterized in the hemocytes of *C. gigas* oyster [87]. Its pattern of expression
614 and its ability to increase the permeability of the cytoplasmic membrane of *Escherichia coli*
615 demonstrated its involvement in the first line of defence against bacterial invasion. Finally,
616 histones were highly represented in zebra mussel hemocytes (1,323 SCs) compared to
617 plasma (58 SCs) (Table 1). Nikapitiya et al. [88] reviewed the role of histone in the immune
618 responses of aquatic invertebrates. In addition to their role in chromatin structure, histones
619 are upregulated during immune challenges and exert antimicrobial peptide activity against
620 bacteria and parasites [89]. Moreover, the release of antimicrobial histones and DNA in
621 plasma by hemocytes of the pacific oyster have been associated with extracellular trap used
622 to counteract pathogen propagation. Thus, the presence of histones in the plasma of the
623 zebra mussel could suggest that extracellular trap mechanisms may also exist in the immune
624 defence of *D. polymorpha*.

625 Many important immune processes are controlled by the collaborative action of
626 proteases and their inhibitors which may act either on endogenous proteins, by cleaving
627 their regulatory subunits and permitting their biological activity, or by causing the
628 inactivation and degradation of proteins produced by microorganisms [33]. Proteases and
629 protease inhibitors known to be involved in the immune system of aquatic invertebrates
630 have been identified in the hemolymph of zebra mussels. Among proteases, numerous

631 cathepsins were observed in both compartments of the hemolymph, four times more
632 abundant in hemocytes than in plasma (Table 1). In bivalves, the endopeptidase cathepsins
633 have been identified in the razor clam (*Sinonovacula constricta*), the pearl oyster (*Pinctada*
634 *fucata*) and the Pacific oyster [90–92]. While the implication of cathepsins in vertebrate
635 immunity is well documented, its role within the innate immune defense of invertebrates is
636 still poorly understood. However, recent investigations showed that gene expression of
637 cathepsins was enhanced after bacterial and viral challenges in the razor clam and the
638 freshwater prawn (*Macrobrachium rosenbergii*), suggesting its implication in the lysis of
639 micro-organisms [91,93]. In vertebrates, cathepsins are involved in the regulation of defence
640 activity and served as a chemical barrier against microbial invasion [94]. Protease inhibitors
641 are used by bivalves to counteract the action of proteases produced by microorganisms [33].
642 Three main types of protease inhibitors potentially involved in the immunity of *D.*
643 *polymorpha* were identified. While serine protease inhibitors were identified exclusively in
644 hemocytes, cystatins and alpha-2-macroglobulins were found in both hemocytes and plasma
645 (Table 1). In the hemolymph of the eastern oyster *C. virginica*, serine protease inhibitors
646 have been implicated in the resistance against the parasite *Perkinsus marinus* and showed
647 an ability to inhibit the major extracellular protease secreted by this pathogen [95]. Cystatin
648 B is a cysteine protease inhibitor, which targets predominantly the cathepsin L-type [96].
649 Like that of cathepsins, the role of cystatins in immunity is not well documented for
650 invertebrates. However, Cystatin B is produced in responses to bacterial infections in the
651 leech (*Theromyzon tessulatum*), the disk abalone (*Haliotis discus discus*) and the Manila clam
652 (*Ruditapes philippinarum*) [97–99]. All authors showed that cystatin B was sensitive to
653 pathological conditions and suggested its implication in the inactivation of proteases
654 produced by invading pathogens. Alpha2-macroglobulin genes have been characterized in
655 the hemocytes of many bivalves, such as the sea scallop (*Chlamys farreri*) and the pearl
656 oyster (*Pinctada fucata*) [100,101]. This protein is designed as a conserved arm of the innate
657 immune system [102]. In invertebrates, this protease inhibitor has been shown to bind and
658 neutralize diverse virulent proteases introduced by microorganisms [103]. Moreover,
659 Ponprateep et al. [104] suggested that alpha2-macroglobulin was able to modulate the pro-
660 phenoloxylase system of the Pacific white shrimp, *Litopenaeus vannamei*.

661

662 3.3.5 The Complement system

663

664 The complement system is a complex molecular machinery of the highest importance in
665 the recognition and elimination of pathogens. Complement system acts as a critical first-line
666 responder of innate immunity, its main functions are to discriminate and eliminate
667 pathogens and promote inflammatory responses [105]. While the complement system has
668 been well characterized in vertebrates, recent investigations confirmed its existence in
669 invertebrates [106–109]. Core component of the complement system (C3 and C2/factor B)
670 and recognition molecules involved in the activation of the complement system have been

671 recently documented in the hemolymph of bivalve species using genomics and
672 transcriptomics analysis [110–112]. Moreover, key components of the complement such as
673 C1q domain containing protein, C3 component and complement factor B were modulated in
674 the hemolymph and extrapalleal fluid of bivalves during infectious experiments [110,112–
675 114].

676 Gerdol et al. [33] discussed on mechanisms of the complement system in bivalves, which
677 is devoid of the membrane attack complex (MAC) and antigen-complexed immunoglobulins.
678 In this system, authors hypothesized that the complex arsenal of soluble PRRs existing in the
679 hemolymph of bivalves, including C1q domain containing protein, CLTs and FReDs may
680 participate in the activation of the complement system. Then, it is not clear how their
681 recognition signals would converge towards C3, since no clear homologue of the serine
682 proteases involved in the cleavage of C3 component has been identified in bivalves [33].
683 According to Wang et al. [112], complement components of oyster probably function as
684 opsonins to assist the phagocytose of pathogens by hemocytes rather than direct lysis of the
685 pathogens. However, pore-forming enzymes such as lysozymes and BPIs could replace the
686 terminal pathway of the complement system, leading to destruction of microbial invaders
687 [33].

688 Numerous proteins related to the complement system were identified in the hemolymph
689 of the zebra mussel. Among them, 9,771 spectra were attributed to the complement
690 component C3 in plasma, representing 16.45% of the proteome, which makes it the most
691 abundant protein identified in plasma (Table 1). C3 component is known to play a pivotal
692 role in the complement system and was considered as the “Swiss Army Knife” of innate
693 immunity and host defence [115]. In addition to the complement C3, proteins containing
694 C1q domains and complement factor B-like proteins were also highly represented in plasma
695 compared to hemocytes (Table 1). Like in vertebrates, proteins containing C1q domains
696 work as PRRs that recognize microorganisms and may trigger the activation of the
697 complement pathways in bivalves [112]. Complement Factor B is a serine protease which is
698 considered as an important regulator of the activation of the component system. Similarly to
699 the complement factor B characterized in the clam *Ruditapes decussatus*, proteins
700 annotated as the complement factor B in the hemolymph of *D. polymorpha* were composed
701 of a serine protease domain and complement control protein modules [107]. In total,
702 proteins related to the complement system represent more than 20% of the plasma
703 proteome (Table1). This suggests a strong involvement of the complement in the first line
704 defence against invaders. However, the mechanisms of action of this process remain
705 uncertain and require further analysis.

706

707 **3.3.6 Proteins involved in apoptotic process**

708

709 The establishment and functioning of the immune system involve apoptosis processes,
710 which actively contribute to the immune system homeostasis. Well-characterized and
711 conserved in vertebrates and invertebrates, this system involves two distinct signalling and
712 execution pathways (intrinsic and extrinsic pathways) that lead to cellular suicide to support
713 the immune system [24]. In molluscs, many studies have highlighted the essential role of
714 hemocyte apoptosis in the immune defence against pathogens, protozoan parasites and
715 environmental pollutants such as herbicides, insecticides and pharmaceutical products.
716 Regulators and effectors of apoptosis were detected mainly in the hemocytes (590 SCs) of
717 *Dreissena polymorpha* compared to plasma (114 SCs) (Table 1). Among them, cytochrome-c
718 (cyto-c), apoptosis regulator BAX, apoptosis inducing factor (AIF) and several inhibitors of
719 apoptosis (IAPs) were identified in the hemocyte compartments. These proteins are well
720 known to be involved in the intrinsic mitochondrial apoptosis pathway. Briefly pro-apoptotic
721 (Bax) and anti-apoptotic (IAP) members of the Bcl-2 family proteins control the
722 mitochondrial membrane permeability [116]. An increase in mitochondrial membrane
723 permeability generates the release of mitochondrial components and proteins into the
724 cytoplasm (cyto-c and AIF) which contributes to the formation of an apoptosome complex
725 (caspase 9/ cyto-c) that activates caspases and triggers apoptosis [116]. Caspase proteins were
726 also identified mainly in hemocytes except for the caspase 3 which was identified in both
727 compartments. According to the literature, this latter is commonly described in intracellular
728 compartments. Thus, caspase 3 may end up in the plasma due to cell lyses caused by the
729 centrifugation of hemocytes. Caspases are a class of cysteine proteases which participate in
730 the initiation and the execution of intrinsic and extrinsic apoptosis pathways [116].
731 Depending on the mechanism of activation and their place in the cascade of events,
732 caspases are divided into two subgroups. In the hemocytes of *D. polymorpha*, initiating
733 caspases (8) and effective caspases (3 and 7) were detected by proteogenomics. Other *D.*
734 *polymorpha* proteins involved in apoptosis were related to MAPKs and NF- κ B immune
735 signalling pathways which are involved in the regulation of both intrinsic and extrinsic
736 apoptosis pathways [116].

737

738 4. Conclusion

739

740 Next-generation proteomics is an interesting tool for a better understanding of the
741 physiological processes of non-model organisms like *D. polymorpha*. In the present study, we
742 applied state-of-the-art proteogenomics to document a large set of proteins from the
743 hemolymph of the zebra mussel. We established the largest repertoire of proteins (3,020
744 identified by tandem mass spectrometry with at least two peptide sequences) for this not-
745 yet genome sequenced sentinel organism. While several proteins were shared between
746 plasma and hemocyte compartments, many of them were detected exclusively in the
747 plasma, suggesting that proteins of plasma could have different origins. The GO analysis
748 performed on immune-related proteins revealed that these two compartments exert closely

749 related immune functions, with few differences. Regarding the immune proteins of
750 hemocytes and plasma, many proteins involved in recognition, internalization and
751 elimination of microorganisms were identified, which underlined the importance of both
752 compartments in the management of foreign particles. The proteogenomics analysis has
753 provided access to identify immune proteins rarely observed in the hemolymph of bivalves.
754 Indeed, a natterin-like protein was observed in high abundance in hemocytes and plasma.
755 While this protein is known to be involved in the recognition and elimination of
756 microorganisms in fish species, its role in bivalve immunity remains to be investigated. Like
757 natterin, a thaumatin-like protein was abundant in the hemolymph of *D. polymorpha* and its
758 role in the immune defence of bivalves is still poorly studied and need further investigations.
759 Furthermore, proteogenomic analysis of the plasma compartment revealed new insights into
760 immune mechanisms occurring in the plasma of bivalves. For example, a large amount of
761 complement-related proteins observed in plasma could confirm the existence of a
762 complement system in bivalves and suggest a strong implication of this mechanism in the
763 management of invaders.

764 Considering this study, dissolved immune proteins from the plasma compartment work
765 in a complementary fashion to hemocytes to neutralize invading organisms and manage
766 immune responses of *D. polymorpha*. Transcriptomic analysis of the hemolymph only
767 focusses on the modulation of hemocyte genes and does not take into consideration
768 complex post-transcriptional regulation and post-translational modifications of polypeptides
769 that occur at large scale in animals. Our proteogenomics strategy proved its capacity to map
770 and quantify the immune proteins in both compartments. Taking into account the possible
771 immune interactions between hemocytes and plasma, such analysis provides a better
772 understanding of immune responses. Thus, we recommend that future research on bivalve
773 immunology should take into consideration the molecular responses from the whole
774 hemolymphatic tissue, including both hemocyte and plasma compartments. Further
775 investigations with targeted quantitative proteomics are now needed to characterize the
776 dynamic of immune proteins during stressful conditions and validate their implications in the
777 management of contaminants [117,118]. This preliminary study paves the way for future
778 proteomics investigations on zebra mussel immunity. This study contributes to the
779 understanding of immune responses of *D. polymorpha* and the discovery of novel molecular
780 immuno-markers that could be used in aquatic ecotoxicology.

781

782 **Acknowledgment**

783 This PhD work was funded by the University of Reims Champagne-Ardenne (project
784 PROMETHEE). Financial support was provided by the French National Research Agency (ANR)
785 project MOBIDIC (Mollusc bivalves as indicators of contamination of water bodies by
786 protozoan parasites, Project ANR- 15-CE34-0005). The authors are grateful to Dr. Antoine LE-
787 GUERNIC, Dr Damien RIOULT and Isabelle BONNARD, technician, for their contribution and
788 advice to this study.

789
790
791

Reference

- 792 [1] B. Allam, D. Raftos, Immune responses to infectious diseases in bivalves, *Journal of*
793 *Invertebrate Pathology*. 131 (2015) 121–136. doi:10.1016/j.jip.2015.05.005.
- 794 [2] A. Binelli, C. Della Torre, S. Magni, M. Parolini, Does zebra mussel (*Dreissena polymorpha*)
795 represent the freshwater counterpart of *Mytilus* in ecotoxicological studies? A critical review,
796 *Environ. Pollut.* 196 (2015) 386–403. doi:10.1016/j.envpol.2014.10.023.
- 797 [3] M.P. Ladeiro, I. Barjhoux, A. Bigot-Clivot, M. Bonnard, E. David, O. Dedourge-Geffard, E.
798 Geba, E. Lance, M. Lepretre, G. Magniez, D. Rioult, D. Aubert, I. Villena, G. Daniele, A.
799 Salvador, E. Vulliet, J. Armengaud, A. Geffard, Mussel as a Tool to Define Continental
800 Watershed Quality, in: S. Ray (Ed.), *Organismal and Molecular Malacology*, InTech, 2017.
801 doi:10.5772/67995.
- 802 [4] W. Xu, M. Faisal, Identification of the molecules involved in zebra mussel (*Dreissena*
803 *polymorpha*) hemocytes host defense, *Comparative Biochemistry and Physiology Part B:*
804 *Biochemistry and Molecular Biology*. 154 (2009) 143–149. doi:10.1016/j.cbpb.2009.05.012.
- 805 [5] M. Gómez-Chiarri, X. Guo, A. Tanguy, Y. He, D. Proestou, The use of -omic tools in the study
806 of disease processes in marine bivalve mollusks, *Journal of Invertebrate Pathology*. 131
807 (2015) 137–154. doi:10.1016/j.jip.2015.05.007.
- 808 [6] J. Armengaud, Microbiology and proteomics, getting the best of both worlds!, *Environ.*
809 *Microbiol.* 15 (2013) 12–23. doi:10.1111/j.1462-2920.2012.02811.x.
- 810 [7] J. Armengaud, J. Trapp, O. Pible, O. Geffard, A. Chaumot, E.M. Hartmann, Non-model
811 organisms, a species endangered by proteogenomics, *Journal of Proteomics*. 105 (2014) 5–
812 18. doi:10.1016/j.jprot.2014.01.007.
- 813 [8] J. Trapp, O. Geffard, G. Imbert, J.-C. Gaillard, A.-H. Davin, A. Chaumot, J. Armengaud,
814 Proteogenomics of *Gammarus fossarum* to Document the Reproductive System of
815 Amphipods, *Mol Cell Proteomics*. 13 (2014) 3612–3625. doi:10.1074/mcp.M114.038851.
- 816 [9] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D.
817 Eccles, B. Li, M. Lieber, M.D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R.
818 Westerman, T. William, C.N. Dewey, R. Henschel, R.D. LeDuc, N. Friedman, A. Regev, De novo
819 transcript sequence reconstruction from RNA-seq using the Trinity platform for reference
820 generation and analysis, *Nature Protocols*. 8 (2013) 1494–1512. doi:10.1038/nprot.2013.084.
- 821 [10] V. Vaibhav, M. Lepretre, E. Thompson, D. Raftos, P.A. Haynes, Biomarkers of Winter
822 Mortality resistance in selectively bred Sydney rock oysters (*Saccostrea glomerata*),
823 *Aquaculture*. 465 (2016) 323–329. doi:10.1016/j.aquaculture.2016.09.006.
- 824 [11] E.M. Hartmann, F. Allain, J.-C. Gaillard, O. Pible, J. Armengaud, Taking the Shortcut for High-
825 Throughput Shotgun Proteomic Analysis of Bacteria, in: A.C. Vergunst, D. O’Callaghan (Eds.),
826 *Host-Bacteria Interactions: Methods and Protocols*, Springer New York, New York, NY, 2014:
827 pp. 275–285. doi:10.1007/978-1-4939-1261-2_16.
- 828 [12] G. Klein, C. Mathé, M. Biola-Clier, S. Devineau, E. Drouineau, E. Hatem, L. Marichal, B. Alonso,
829 J.-C. Gaillard, G. Lagniel, J. Armengaud, M. Carrière, S. Chédin, Y. Boulard, S. Pin, J.-P. Renault,
830 J.-C. Aude, J. Labarre, RNA-binding proteins are a major target of silica nanoparticles in cell
831 extracts, *Nanotoxicology*. 10 (2016) 1555–1564. doi:10.1080/17435390.2016.1244299.
- 832 [13] J.A. Christie-Oleza, B. Fernandez, B. Nogales, R. Bosch, J. Armengaud, Proteomic insights into
833 the lifestyle of an environmentally relevant marine bacterium, *The ISME Journal*. 6 (2012)
834 124–135. doi:10.1038/ismej.2011.86.
- 835 [14] A. Marchler-Bauer, S. Lu, J.B. Anderson, F. Chitsaz, M.K. Derbyshire, C. DeWeese-Scott, J.H.
836 Fong, L.Y. Geer, R.C. Geer, N.R. Gonzales, M. Gwadz, D.I. Hurwitz, J.D. Jackson, Z. Ke, C.J.
837 Lanczycki, F. Lu, G.H. Marchler, M. Mulkolokandov, M.V. Omelchenko, C.L. Robertson, J.S. Song,
838 N. Thanki, R.A. Yamashita, D. Zhang, N. Zhang, C. Zheng, S.H. Bryant, CDD: a Conserved

- 839 Domain Database for the functional annotation of proteins, *Nucleic Acids Research*. 39
840 (2011) D225–D229. doi:10.1093/nar/gkq1189.
- 841 [15] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D.J. Kundu, A.
842 Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg,
843 S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A.F. Jarnuczak, T. Ternent, A. Brazma, J.A.
844 Vizcaíno, The PRIDE database and related tools and resources in 2019: improving support for
845 quantification data, *Nucleic Acids Res.* 47 (2019) D442–D450. doi:10.1093/nar/gky1106.
- 846 [16] P. Jagtap, J. Goslinga, J.A. Kooren, T. McGowan, M.S. Wroblewski, S.L. Seymour, T.J. Griffin, A
847 two-step database search method improves sensitivity in peptide sequence matches for
848 metaproteomics and proteogenomics studies, *Proteomics*. 13 (2013) 1352–1357.
849 doi:10.1002/pmic.201200352.
- 850 [17] J. Trapp, C. Almunia, J.C. Gaillard, O. Pible, A. Chaumot, O. Geffard, J. Armengaud,
851 Proteogenomic insights into the core-proteome of female reproductive tissues from
852 crustacean amphipods., *J Proteomics*. 135 (2016) 51–61. doi:10.1016/j.jprot.2015.06.017.
- 853 [18] A. Campos, I. Apraiz, R.R. da Fonseca, S. Cristobal, Shotgun analysis of the marine mussel
854 *Mytilus edulis* hemolymph proteome and mapping the innate immunity elements,
855 *PROTEOMICS*. 15 (2015) 4021–4029. doi:10.1002/pmic.201500118.
- 856 [19] K.M. Halanych, K.M. Kocot, Genome evolution: Shellfish genes, *Nature Ecology & Evolution*. 1
857 (2017) 0142. doi:10.1038/s41559-017-0142.
- 858 [20] S. Sricharoen, J.J. Kim, S. Tunkijjanukij, I. Söderhäll, Exocytosis and proteomic analysis of the
859 vesicle content of granular hemocytes from a crayfish, *Developmental & Comparative*
860 *Immunology*. 29 (2005) 1017–1031. doi:10.1016/j.dci.2005.03.010.
- 861 [21] L. Canesi, M. Betti, C. Ciacci, L.C. Lorusso, C. Pruzzo, G. Gallo, Cell signalling in the immune
862 response of mussel hemocytes, *Invertebrate Survival Journal*. 3 (2006) 40–49.
- 863 [22] F. Qu, Z. Xiang, J. Li, S. Xiao, F. Mao, Y. Qin, Y. Zhou, H. Ma, Z. Yu, A molluscan extracellular
864 signal-regulated kinase is involved in host response to immune challenges in vivo and in vitro,
865 *Fish & Shellfish Immunology*. 62 (2017) 311–319. doi:10.1016/j.fsi.2017.01.038.
- 866 [23] F. Le Foll, D. Rioult, S. Boussa, J. Pasquier, Z. Dagher, F. Leboulenger, Characterisation of
867 *Mytilus edulis* hemocyte subpopulations by single cell time-lapse motility imaging, *Fish &*
868 *Shellfish Immunology*. 28 (2010) 372–386. doi:10.1016/j.fsi.2009.11.011.
- 869 [24] I.M. Sokolova, Apoptosis in molluscan immune defense, *Invertebrate Survival Journal*. 6
870 (2009) 49–58.
- 871 [25] T. Balbi, K. Cortese, C. Ciacci, G. Bellese, L. Vezzulli, C. Pruzzo, L. Canesi, Autophagic processes
872 in *Mytilus galloprovincialis* hemocytes: Effects of *Vibrio tapetis*, *Fish & Shellfish Immunology*.
873 73 (2018) 66–74. doi:10.1016/j.fsi.2017.12.003.
- 874 [26] D. Malagoli, E. Ottaviani, Cross-talk among immune and neuroendocrine systems in molluscs
875 and other invertebrate models, *Hormones and Behavior*. 88 (2017) 41–44.
876 doi:10.1016/j.yhbeh.2016.10.015.
- 877 [27] L. Cerenius, S. Kawabata, B.L. Lee, M. Nonaka, K. Söderhäll, Proteolytic cascades and their
878 involvement in invertebrate immunity, *Trends in Biochemical Sciences*. 35 (2010) 575–583.
879 doi:10.1016/j.tibs.2010.04.006.
- 880 [28] P. Venier, L. Varotto, U. Rosani, C. Millino, B. Celegato, F. Bernante, G. Lanfranchi, B. Novoa,
881 P. Roch, A. Figueras, others, Insights into the innate immunity of the Mediterranean mussel
882 *Mytilus galloprovincialis*, *BMC Genomics*. 12 (2011) 1.
- 883 [29] G. Magalhaes, M. Lopesferreira, I. Junqueira de Azevedo, P. Spencer, M. Araujo, F. Portaro, L.
884 Ma, R. Valente, L. Juliano, J. Fox, Natterins, a new class of proteins with kininogenase activity
885 characterized from fish venom, *Biochimie*. 87 (2005) 687–699.
886 doi:10.1016/j.biochi.2005.03.016.
- 887 [30] E.N. Komegae, A.D. Ramos, A.K. Oliveira, S.M. de T. Serrano, M. Lopes-Ferreira, C. Lima,
888 Insights into the local pathogenesis induced by fish toxins: role of natterins and natterectin in
889 the disruption of cell-cell and cell-extracellular matrix interactions and modulation of cell
890 migration, *Toxicon*. 58 (2011) 509–517. doi:10.1016/j.toxicon.2011.08.012.

- 891 [31] B. Rajan, D. Patel, Y. Kitani, K. Viswanath, M. Brinchmann, Novel mannose binding natterin-
892 like protein in the skin mucus of Atlantic cod (*Gadus morhua*), *Fish & Shellfish Immunology*.
893 68 (2017). doi:10.1016/j.fsi.2017.07.039.
- 894 [32] Z. Xue, X. Liu, Y. Pang, T. Yu, R. Xiao, M. Jin, Y. Han, P. Su, J. Wang, L. Lv, F. Wu, Q. Li,
895 Characterization, phylogenetic analysis and cDNA cloning of natterin-like gene from the
896 blood of lamprey, *Lampetra japonica*, *Immunology Letters*. 148 (2012) 1–10.
897 doi:10.1016/j.imlet.2012.08.005.
- 898 [33] M. Gerdol, M. Gomez-Chiarri, M.G. Castillo, A. Figueras, G. Fiorito, R. Moreira, B. Novoa, A.
899 Pallavicini, G. Ponte, K. Roumbedakis, P. Venier, G.R. Vasta, *Immunity in Molluscs:
900 Recognition and Effector Mechanisms, with a Focus on Bivalvia*, in: E.L. Cooper (Ed.),
901 *Advances in Comparative Immunology*, Springer International Publishing, Cham, 2018: pp.
902 225–341. doi:10.1007/978-3-319-76768-0_11.
- 903 [34] X. Jing, E.P. Espinosa, M. Perrigault, B. Allam, Identification, molecular characterization and
904 expression analysis of a mucosal C-type lectin in the eastern oyster, *Crassostrea virginica*,
905 *Fish & Shellfish Immunology*. 30 (2011) 851–858. doi:10.1016/j.fsi.2011.01.007.
- 906 [35] M. Huang, H. Zhang, S. Jiang, L. Wang, R. Liu, Q. Yi, L. Song, An EPD/WSD motifs containing C-
907 type lectin from *Argopectens irradians* recognizes and binds microbes with broad spectrum,
908 *Fish & Shellfish Immunology*. 43 (2015) 287–293. doi:10.1016/j.fsi.2014.12.035.
- 909 [36] J. Yang, M. Huang, H. Zhang, L. Wang, H. Wang, L. Wang, L. Qiu, L. Song, CfLec-3 from scallop:
910 an entrance to non-self recognition mechanism of invertebrate C-type lectin, *Scientific
911 Reports*. 5 (2015) 10068. doi:10.1038/srep10068.
- 912 [37] Y. Lu, H. Zhang, D. Cheng, H. Liu, S. Li, H. Ma, H. Zheng, A multi-CRD C-type lectin gene Cnlec-
913 1 enhance the immunity response in noble scallop *Chlamys nobilis* with higher carotenoids
914 contents through up-regulating under different immunostimulants, *Fish & Shellfish
915 Immunology*. 83 (2018) 37–44. doi:10.1016/j.fsi.2018.09.014.
- 916 [38] M. Gerdol, P. Venier, An updated molecular basis for mussel immunity, *Fish & Shellfish
917 Immunology*. 46 (2015) 17–38. doi:10.1016/j.fsi.2015.02.013.
- 918 [39] I. Hasan, M. Gerdol, Y. Fujii, S. Rajia, Y. Koide, D. Yamamoto, S.M.A. Kawsar, Y. Ozeki, cDNA
919 and Gene Structure of MytiLec-1, A Bacteriostatic R-Type Lectin from the Mediterranean
920 Mussel (*Mytilus galloprovincialis*), *Mar Drugs*. 14 (2016). doi:10.3390/md14050092.
- 921 [40] G.R. Vasta, Galectins as Pattern Recognition Receptors: Structure, Function, and Evolution,
922 *Adv Exp Med Biol*. 946 (2012) 21–36. doi:10.1007/978-1-4614-0106-3_2.
- 923 [41] G.R. Vasta, C. Feng, M.A. Bianchet, T.R. Bachvaroff, S. Tasumi, Structural, functional, and
924 evolutionary aspects of galectins in aquatic mollusks: From a sweet tooth to the Trojan
925 horse, *Fish & Shellfish Immunology*. 46 (2015) 94–106. doi:10.1016/j.fsi.2015.05.012.
- 926 [42] Z. Bai, L. Zhao, X. Chen, Q. Li, J. Li, A galectin contributes to the innate immune recognition
927 and elimination of pathogens in the freshwater mussel *Hyriopsis cumingii*, *Developmental &
928 Comparative Immunology*. 73 (2017) 36–45. doi:10.1016/j.dci.2017.03.008.
- 929 [43] P.C. Hanington, S.-M. Zhang, The Primary Role of Fibrinogen-Related Proteins in
930 Invertebrates Is Defense, Not Coagulation, *J Innate Immun*. 3 (2010) 17–27.
931 doi:10.1159/000321882.
- 932 [44] H. Zhang, L. Wang, L. Song, X. Song, B. Wang, C. Mu, Y. Zhang, A fibrinogen-related protein
933 from bay scallop *Argopecten irradians* involved in innate immunity as pattern recognition
934 receptor, *Fish Shellfish Immunol*. 26 (2009) 56–64. doi:10.1016/j.fsi.2008.07.019.
- 935 [45] K.-Y. Zhang, W.-J. Yuan, J.-D. Xu, J.-X. Wang, Cation-dependent mannose-6-phosphate
936 receptor functions as a pattern recognition receptor in anti-bacterial immunity of
937 *Marsupenaeus japonicus*, *Developmental & Comparative Immunology*. 89 (2018) 122–130.
938 doi:10.1016/j.dci.2018.08.011.
- 939 [46] J. Yang, L. Qiu, L. Wang, X. Wei, H. Zhang, Y. Zhang, L. Liu, L. Song, CfLGBP, a pattern
940 recognition receptor in *Chlamys farreri* involved in the immune response against various
941 bacteria, *Fish & Shellfish Immunology*. 29 (2010) 825–831. doi:10.1016/j.fsi.2010.07.025.

- 942 [47] N. Itoh, K.G. Takahashi, A novel peptidoglycan recognition protein containing a goose-type
943 lysozyme domain from the Pacific oyster, *Crassostrea gigas*, *Molecular Immunology*. 46
944 (2009) 1768–1774. doi:10.1016/j.molimm.2009.01.022.
- 945 [48] D. Ni, L. Song, L. Wu, Y. Chang, Y. Yu, L. Qiu, L. Wang, Molecular cloning and mRNA
946 expression of peptidoglycan recognition protein (PGRP) gene in bay scallop (*Argopecten*
947 *irradians*, Lamarck 1819), *Developmental & Comparative Immunology*. 31 (2007) 548–558.
948 doi:10.1016/j.dci.2006.09.001.
- 949 [49] F. Li, S. Huang, L. Wang, J. Yang, H. Zhang, L. Qiu, L. Li, L. Song, A macrophage migration
950 inhibitory factor like gene from scallop *Chlamys farreri*: Involvement in immune response and
951 wound healing, *Developmental & Comparative Immunology*. 35 (2011) 62–71.
952 doi:10.1016/j.dci.2010.08.009.
- 953 [50] D. Wang, D. Yang, Q. Wang, Y. Zhao, C. Li, Q. Wei, Y. Han, J. Zhao, Two macrophage migration
954 inhibitory factors (MIFs) from the clam *Ruditapes philippinarum*: Molecular characterization,
955 localization and enzymatic activities, *Fish & Shellfish Immunology*. 78 (2018) 158–168.
956 doi:10.1016/j.fsi.2018.04.026.
- 957 [51] L. Martín-Gómez, A. Villalba, M.J. Carballal, E. Abollo, Molecular characterisation of TNF, AIF,
958 dermatopontin and VAMP genes of the flat oyster *Ostrea edulis* and analysis of their
959 modulation by diseases, *Gene*. 533 (2014) 208–217. doi:10.1016/j.gene.2013.09.085.
- 960 [52] Y. Zhang, J. Li, F. Yu, X. He, Z. Yu, Allograft inflammatory factor-1 stimulates hemocyte
961 immune activation by enhancing phagocytosis and expression of inflammatory cytokines in
962 *Crassostrea gigas*, *Fish & Shellfish Immunology*. 34 (2013) 1071–1077.
963 doi:10.1016/j.fsi.2013.01.014.
- 964 [53] J. Li, J. Chen, Y. Zhang, Z. Yu, Expression of allograft inflammatory factor-1 (AIF-1) in response
965 to bacterial challenge and tissue injury in the pearl oyster, *Pinctada martensii*, *Fish & Shellfish*
966 *Immunology*. 34 (2013) 365–371. doi:10.1016/j.fsi.2012.11.012.
- 967 [54] C. Dong, R.J. Davis, R.A. Flavell, MAP Kinases in the Immune Response, *Annual Review of*
968 *Immunology*. 20 (2002) 55–72. doi:10.1146/annurev.immunol.20.091301.131133.
- 969 [55] H. Shi, X. Yan, X. Xu, L. Ruan, Molecular cloning and characterization of a cDNA encoding
970 extracellular signal-regulated kinase from *Litopenaeus vannamei*, *Fish & Shellfish*
971 *Immunology*. 33 (2012) 813–820. doi:10.1016/j.fsi.2012.07.008.
- 972 [56] Y. Sun, L. Zhang, M. Zhang, R. Li, Y. Li, X. Hu, S. Wang, Z. Bao, Characterization of three
973 mitogen-activated protein kinases (MAPK) genes reveals involvement of ERK and JNK, not
974 p38 in defense against bacterial infection in Yesso scallop *Patinopecten yessoensis*, *Fish &*
975 *Shellfish Immunology*. 54 (2016) 507–515. doi:10.1016/j.fsi.2016.04.139.
- 976 [57] R.E. Randall, S. Goodbourn, Interferons and viruses: an interplay between induction,
977 signalling, antiviral responses and virus countermeasures, *Journal of General Virology*. 89
978 (2008) 1–47. doi:10.1099/vir.0.83391-0.
- 979 [58] T.J. Green, D. Raftos, P. Speck, C. Montagnani, Antiviral immunity in marine molluscs, *Journal*
980 *of General Virology*. 96 (2015) 2471–2482. doi:10.1099/jgv.0.000244.
- 981 [59] L. Canesi, M. Betti, C. Ciacci, B. Citterio, C. Pruzzo, G. Gallo, Tyrosine kinase-mediated cell
982 signalling in the activation of *Mytilus* hemocytes: possible role of STAT-like proteins, *Biology*
983 *of the Cell*. 95 (2003) 603–613. doi:10.1016/j.biolcel.2003.09.006.
- 984 [60] T.J. Green, C. Montagnani, Poly I:C induces a protective antiviral immune response in the
985 Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus
986 (OsHV-1 μ var), *Fish & Shellfish Immunology*. 35 (2013) 382–388.
987 doi:10.1016/j.fsi.2013.04.051.
- 988 [61] T. Renault, N. Faury, V. Barbosa-Solomieu, K. Moreau, Suppression subtractive hybridisation
989 (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster,
990 *Crassostrea gigas*, challenged with Ostreid herpesvirus 1, *Developmental & Comparative*
991 *Immunology*. 35 (2011) 725–735. doi:10.1016/j.dci.2011.02.004.
- 992 [62] S. Koutsogiannaki, M. Kaloyianni, Signaling molecules involved in immune responses in
993 mussels, *Inv Surv J*. 7 (2010) 11–21.

- 994 [63] M.-O. Fauvarque, M.J. Williams, *Drosophila* cellular immunity: a story of migration and
995 adhesion, *J Cell Sci.* 124 (2011) 1373–1382. doi:10.1242/jcs.064592.
- 996 [64] P. Rougerie, V. Miskolci, D. Cox, Generation of membrane structures during phagocytosis and
997 chemotaxis of macrophages: role and regulation of the actin cytoskeleton, *Immunol Rev.* 256
998 (2013) 222–239. doi:10.1111/imr.12118.
- 1000 [65] A.J. Ridley, M.A. Schwartz, K. Burridge, R.A. Firtel, M.H. Ginsberg, G. Borisy, J.T. Parsons, A.R.
1001 Horwitz, Cell Migration: Integrating Signals from Front to Back, *Science.* 302 (2003) 1704–
1709. doi:10.1126/science.1092053.
- 1002 [66] C. Rosales, E. Uribe-Querol, Phagocytosis: A Fundamental Process in Immunity, *BioMed*
1003 *Research International.* (2017). doi:10.1155/2017/9042851.
- 1004 [67] Y.-T. Lau, L. Gambino, B. Santos, E. Pales Espinosa, B. Allam, Regulation of oyster (*Crassostrea*
1005 *virginica*) hemocyte motility by the intracellular parasite *Perkinsus marinus*: A possible
1006 mechanism for host infection, *Fish & Shellfish Immunology.* 78 (2018) 18–25.
1007 doi:10.1016/j.fsi.2018.04.019.
- 1008 [68] Y. Zhang, H. Wang, Integrin signalling and function in immune cells, *Immunology.* 135 (2012)
1009 268–275. doi:10.1111/j.1365-2567.2011.03549.x.
- 1010 [69] C.G.A. Moreira, A. Jacinto, S. Prag, *Drosophila* integrin adhesion complexes are essential for
1011 hemocyte migration in vivo, *Biol Open.* 2 (2013) 795–801. doi:10.1242/bio.20134564.
- 1012 [70] L. Ballarin, M. Scanferla, F. Cima, A. Sabbadin, Phagocyte spreading and phagocytosis in the
1013 compound ascidian *Botryllus schlosseri*: evidence for an integrin-like, RGD-dependent
1014 recognition mechanism, *Developmental & Comparative Immunology.* 26 (2002) 345–354.
1015 doi:10.1016/S0145-305X(01)00082-9.
- 1016 [71] K. Terahara, K.G. Takahashi, A. Nakamura, M. Osada, M. Yoda, T. Hiroi, M. Hirasawa, K. Mori,
1017 Differences in integrin-dependent phagocytosis among three hemocyte subpopulations of
1018 the Pacific oyster “*Crassostrea gigas*,” *Developmental & Comparative Immunology.* 30 (2006)
1019 667–683. doi:10.1016/j.dci.2005.09.009.
- 1020 [72] Z. Jia, T. Zhang, S. Jiang, M. Wang, Q. Cheng, M. Sun, L. Wang, L. Song, An integrin from
1021 oyster *Crassostrea gigas* mediates the phagocytosis toward *Vibrio splendidus* through LPS
1022 binding activity, *Developmental & Comparative Immunology.* 53 (2015) 253–264.
1023 doi:10.1016/j.dci.2015.07.014.
- 1024 [73] H.T. McMahon, E. Boucrot, Molecular mechanism and physiological functions of clathrin-
1025 mediated endocytosis, *Nature Reviews Molecular Cell Biology.* 12 (2011) 517–533.
1026 doi:10.1038/nrm3151.
- 1027 [74] L. DeTulleo, T. Kirchhausen, The clathrin endocytic pathway in viral infection, *The EMBO*
1028 *Journal.* 17 (1998) 4585–4593. doi:10.1093/emboj/17.16.4585.
- 1029 [75] P. Cossart, A. Helenius, Endocytosis of Viruses and Bacteria, *Cold Spring Harb Perspect Biol.* 6
1030 (2014). doi:10.1101/cshperspect.a016972.
- 1031 [76] M.-C. Yang, X.-Z. Shi, H.-T. Yang, J.-J. Sun, L. Xu, X.-W. Wang, X.-F. Zhao, J.-X. Wang, Scavenger
1032 Receptor C Mediates Phagocytosis of White Spot Syndrome Virus and Restricts Virus
1033 Proliferation in Shrimp, *PLoS Pathog.* 12 (2016). doi:10.1371/journal.ppat.1006127.
- 1034 [77] T. Jatuyosporn, P. Supungul, A. Tassanakajon, K. Krusong, The essential role of clathrin-
1035 mediated endocytosis in yellow head virus propagation in the black tiger shrimp *Penaeus*
1036 *monodon*, *Developmental & Comparative Immunology.* 44 (2014) 100–110.
1037 doi:10.1016/j.dci.2013.11.017.
- 1038 [78] J.-J. Liu, R. Sturrock, A.K.M. Ekramoddoullah, The superfamily of thaumatin-like proteins: its
1039 origin, evolution, and expression towards biological function, *Plant Cell Rep.* 29 (2010) 419–
1040 436. doi:10.1007/s00299-010-0826-8.
- 1041 [79] Z. Bai, Y. Yin, S. Hu, G. Wang, X. Zhang, J. Li, Identification of Genes Involved in Immune
1042 Response, Microsatellite, and SNP Markers from Expressed Sequence Tags Generated from
1043 Hemocytes of Freshwater Pearl Mussel (*Hyriopsis cumingii*), *Marine Biotechnology.* 11 (2009)
1044 520–530. doi:10.1007/s10126-008-9163-0.

- 1045 [80] E. Pales Espinosa, A. Koller, B. Allam, Proteomic characterization of mucosal secretions in the
1046 eastern oyster, *Crassostrea virginica*, *Journal of Proteomics*. 132 (2016) 63–76.
1047 doi:10.1016/j.jprot.2015.11.018.
- 1048 [81] T. Ganz, Defensins: antimicrobial peptides of innate immunity, *Nature Reviews Immunology*.
1049 3 (2003) 710–720. doi:10.1038/nri1180.
- 1050 [82] R.D. Rosa, A. Santini, J. Fievet, P. Bulet, D. Destoumieux-Garzón, E. Bachère, Big defensins, a
1051 diverse family of antimicrobial peptides that follows different patterns of expression in
1052 hemocytes of the oyster *Crassostrea gigas*, *PLoS ONE*. 6 (2011) e25594.
1053 doi:10.1371/journal.pone.0025594.
- 1054 [83] J. Zhao, L. Song, C. Li, D. Ni, L. Wu, L. Zhu, H. Wang, W. Xu, Molecular cloning, expression of a
1055 big defensin gene from bay scallop *Argopecten irradians* and the antimicrobial activity of its
1056 recombinant protein, *Mol. Immunol.* 44 (2007) 360–368. doi:10.1016/j.molimm.2006.02.025.
- 1057 [84] W. Xu, M. Faisal, Defensin of the zebra mussel (*Dreissena polymorpha*): molecular structure,
1058 in vitro expression, antimicrobial activity, and potential functions, *Mol. Immunol.* 47 (2010)
1059 2138–2147. doi:10.1016/j.molimm.2010.01.025.
- 1060 [85] J.M. Van Herreweghe, C.W. Michiels, Invertebrate lysozymes: Diversity and distribution,
1061 molecular mechanism and in vivo function, *J Biosci.* 37 (2012) 327–348. doi:10.1007/s12038-
1062 012-9201-y.
- 1063 [86] Q. Wang, C. Wang, C. Mu, H. Wu, L. Zhang, J. Zhao, A Novel C-Type Lysozyme from *Mytilus*
1064 *galloprovincialis*: Insight into Innate Immunity and Molecular Evolution of Invertebrate C-
1065 Type Lysozymes, *PLOS ONE*. 8 (2013) e67469. doi:10.1371/journal.pone.0067469.
- 1066 [87] M. Gonzalez, Y. Gueguen, D. Destoumieux-Garzon, B. Romestand, J. Fievet, M. Pugnieri, F.
1067 Roquet, J.-M. Escoubas, F. Vandenbulcke, O. Levy, L. Saune, P. Bulet, E. Bachere, Evidence of
1068 a bactericidal permeability increasing protein in an invertebrate, the *Crassostrea gigas* Cg-
1069 BPI, *Proceedings of the National Academy of Sciences*. 104 (2007) 17759–17764.
1070 doi:10.1073/pnas.0702281104.
- 1071 [88] C. Nikapitiya, T. Dorrington, M. Gomez-Chiarri, The role of histones in the immune responses
1072 of aquatic invertebrates, *ISJ*. 10 (2013) 94–101.
- 1073 [89] J.-K. Seo, J. Stephenson, J.M. Crawford, K.L. Stone, E.J. Noga, American Oyster, *Crassostrea*
1074 *virginica*, Expresses a Potent Antibacterial Histone H2B Protein, *Mar Biotechnol.* 12 (2010)
1075 543–551. doi:10.1007/s10126-009-9240-z.
- 1076 [90] J. Ma, D. Zhang, J. Jiang, S. Cui, H. Pu, S. Jiang, Molecular characterization and expression
1077 analysis of cathepsin L1 cysteine protease from pearl oyster *Pinctada fucata*, *Fish Shellfish*
1078 *Immunol.* 29 (2010) 501–507. doi:10.1016/j.fsi.2010.05.006.
- 1079 [91] D. Niu, K. Jin, L. Wang, B. Feng, J. Li, Molecular characterization and expression analysis of
1080 four cathepsin L genes in the razor clam, *Sinonovacula constricta*, *Fish Shellfish Immunol.* 35
1081 (2013) 581–588. doi:10.1016/j.fsi.2013.06.001.
- 1082 [92] G. Zhang, X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H.
1083 Que, Y. Xie, P.W.H. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang,
1084 J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T.
1085 Domazet-Lošo, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu,
1086 T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, J. Wang, M. Chen, Y. He, F. Tan, X. Song, Q.
1087 Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P.M. Gaffney, S. Wang, L. Luo, Z. She, Y.
1088 Ming, W. Huang, S. Zhang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, J. Wang, Q. Wang, C.E.W.
1089 Steinberg, H. Wang, N. Li, L. Qian, G. Zhang, Y. Li, H. Yang, X. Liu, J. Wang, Y. Yin, J. Wang, The
1090 oyster genome reveals stress adaptation and complexity of shell formation, *Nature*. 490
1091 (2012) 49. doi:10.1038/nature11413.
- 1092 [93] J. Arockiaraj, A.J. Gnanam, D. Muthukrishnan, M.K. Thirumalai, M. Pasupuleti, J. Milton, M.
1093 Kasi, *Macrobrachium rosenbergii* cathepsin L: Molecular characterization and gene
1094 expression in response to viral and bacterial infections, *Microbiological Research*. 168 (2013)
1095 569–579. doi:10.1016/j.micres.2013.04.007.

- 1096 [94] S. Conus, H.-U. Simon, Cathepsins and their involvement in immune responses, *Swiss Med*
1097 *Wkly.* 140 (2010) w13042. doi:10.4414/smw.2010.13042.
- 1098 [95] J.F. La Peyre, Q.-G. Xue, N. Itoh, Y. Li, R.K. Cooper, Serine protease inhibitor cvSI-1 potential
1099 role in the eastern oyster host defense against the protozoan parasite *Perkinsus marinus*,
1100 *Dev. Comp. Immunol.* 34 (2010) 84–92. doi:10.1016/j.dci.2009.08.007.
- 1101 [96] C. Lefebvre, F. Vandenbulcke, B. Bocquet, A. Tasiemski, A. Desmons, M. Verstraete, M. Salzet,
1102 C. Cocquerelle, Cathepsin L and cystatin B gene expression discriminates immune cœlomic
1103 cells in the leech *Theromyzon tessulatum*, *Dev Comp Immunol.* 32 (2008) 795–807.
1104 doi:10.1016/j.dci.2007.11.007.
- 1105 [97] C. Lefebvre, C. Cocquerelle, F. Vandenbulcke, D. Hot, L. Huot, Y. Lemoine, M. Salzet,
1106 Transcriptomic analysis in the leech *Theromyzon tessulatum*: involvement of cystatin B in
1107 innate immunity., *Biochem J.* 380 (2004) 617–625. doi:10.1042/BJ20040478.
- 1108 [98] H.K.A. Premachandra, D.A.S. Elvitigala, I. Whang, E. Kim, M. De Zoysa, B.-S. Lim, S.-Y. Yeo, S.
1109 Kim, M.-A. Park, H.-C. Park, J. Lee, Expression profile of cystatin B ortholog from Manila clam
1110 (*Ruditapes philippinarum*) in host pathology with respect to its structural and functional
1111 properties, *Fish & Shellfish Immunology.* 34 (2013) 1505–1513.
1112 doi:10.1016/j.fsi.2013.03.349.
- 1113 [99] H.K.A. Premachandra, Q. Wan, D.A.S. Elvitigala, M. De Zoysa, C.Y. Choi, I. Whang, J. Lee,
1114 Genomic characterization and expression profiles upon bacterial infection of a novel cystatin
1115 B homologue from disk abalone (*Haliotis discus discus*), *Developmental & Comparative*
1116 *Immunology.* 38 (2012) 495–504. doi:10.1016/j.dci.2012.06.010.
- 1117 [100] H. Ma, K. Mai, W. Xu, Z. Liufu, Molecular cloning of alpha2- macroglobulin in sea scallop
1118 *Chlamys farreri* (Bivalvia, Mollusca), *Fish & Shellfish Immunology.* 18 (2005) 345–349.
1119 doi:10.1016/j.fsi.2004.08.006.
- 1120 [101] Z. Wang, B. Wang, G. Chen, Y. Lu, J. Jian, Z. Wu, An alpha-2 macroglobulin in the pearl oyster
1121 *Pinctada fucata*: Characterization and function in hemocyte phagocytosis of *Vibrio*
1122 *alginolyticus*, *Fish & Shellfish Immunology.* 55 (2016) 585–594. doi:10.1016/j.fsi.2016.06.037.
- 1123 [102] P.B. Armstrong, J.P. Quigley, α 2-macroglobulin: an evolutionarily conserved arm of the
1124 innate immune system, *Developmental & Comparative Immunology.* 23 (1999) 375–390.
- 1125 [103] P.B. Armstrong, Role of α 2-macroglobulin in the immune responses of invertebrates.,
1126 *Invertebrate Survival Journal.* 7 (2010).
- 1127 [104] S. Ponprateep, T. Vatanavicharn, C.F. Lo, A. Tassanakajon, V. Rimphanitchayakit, Alpha-2-
1128 macroglobulin is a modulator of prophenoloxidase system in pacific white shrimp
1129 *Litopenaeus vannamei*, *Fish & Shellfish Immunology.* 62 (2017) 68–74.
1130 doi:10.1016/j.fsi.2016.12.028.
- 1131 [105] D. Ricklin, G. Hajishengallis, K. Yang, J.D. Lambris, Complement: a key system for immune
1132 surveillance and homeostasis, *Nature Immunology.* 11 (2010) 785. doi:10.1038/ni.1923.
- 1133 [106] M.G. Castillo, M.S. Goodson, M. McFall-Ngai, Identification and molecular characterization of
1134 a complement C3 molecule in a lophotrochozoan, the Hawaiian bobtail squid *Euprymna*
1135 *scolopes*, *Developmental & Comparative Immunology.* 33 (2009) 69–76.
1136 doi:10.1016/j.dci.2008.07.013.
- 1137 [107] M. Prado-Alvarez, J. Rotllant, C. Gestal, B. Novoa, A. Figueras, Characterization of a C3 and a
1138 factor B-like in the carpet-shell clam, *Ruditapes decussatus*, *Fish & Shellfish Immunology.* 26
1139 (2009) 305–315. doi:10.1016/j.fsi.2008.11.015.
- 1140 [108] T. Xu, J. Xie, J. Li, M. Luo, S. Ye, X. Wu, Identification of expressed genes in cDNA library of
1141 hemocytes from the RLO-challenged oyster, *Crassostrea ariakensis* Gould with special
1142 functional implication of three complement-related fragments (CaC1q1, CaC1q2 and CaC3),
1143 *Fish & Shellfish Immunology.* 32 (2012) 1106–1116. doi:10.1016/j.fsi.2012.03.012.
- 1144 [109] H. Zhang, L. Song, C. Li, J. Zhao, H. Wang, Q. Gao, W. Xu, Molecular cloning and
1145 characterization of a thioester-containing protein from Zhikong scallop *Chlamys farreri*,
1146 *Molecular Immunology.* 44 (2007) 3492–3500. doi:10.1016/j.molimm.2007.03.008.

- 1147 [110] Y. Chen, K. Xu, J. Li, X. Wang, Y. Ye, P. Qi, Molecular characterization of complement
 1148 component 3 (C3) in *Mytilus coruscus* improves our understanding of bivalve complement
 1149 system, *Fish & Shellfish Immunology*. 76 (2018) 41–47. doi:10.1016/j.fsi.2018.02.044.
 1150 [111] M. Peng, D. Niu, Z. Chen, T. Lan, Z. Dong, T.-N. Tran, J. Li, Expression of a novel complement
 1151 C3 gene in the razor clam *Sinonovacula constricta* and its role in innate immune response
 1152 and hemolysis, *Developmental & Comparative Immunology*. 73 (2017) 184–192.
 1153 doi:10.1016/j.dci.2017.03.027.
 1154 [112] L. Wang, H. Zhang, L. Wang, D. Zhang, Z. Lv, Z. Liu, W. Wang, Z. Zhou, L. Qiu, H. Wang, J. Li, L.
 1155 Song, The RNA-seq analysis suggests a potential multi-component complement system in
 1156 oyster *Crassostrea gigas*, *Developmental & Comparative Immunology*. 76 (2017) 209–219.
 1157 doi:10.1016/j.dci.2017.06.009.
 1158 [113] B. Allam, E. Pales Espinosa, A. Tanguy, F. Jeffroy, C. Le Bris, C. Paillard, Transcriptional
 1159 changes in Manila clam (*Ruditapes philippinarum*) in response to Brown Ring Disease, *Fish &
 1160 Shellfish Immunology*. 41 (2014) 2–11. doi:10.1016/j.fsi.2014.05.022.
 1161 [114] M. Peng, D. Niu, F. Wang, Z. Chen, J. Li, Complement C3 gene: Expression characterization
 1162 and innate immune response in razor clam *Sinonovacula constricta*, *Fish & Shellfish
 1163 Immunology*. 55 (2016) 223–232. doi:10.1016/j.fsi.2016.05.024.
 1164 [115] D. Ricklin, E.S. Reis, D.C. Mastellos, P. Gros, J.D. Lambris, Complement component C3 – The
 1165 “Swiss Army Knife” of innate immunity and host defense, *Immunol Rev*. 274 (2016) 33–58.
 1166 doi:10.1111/imr.12500.
 1167 [116] T. Kiss, Apoptosis and its functional significance in molluscs, *Apoptosis*. 15 (2010) 313–321.
 1168 doi:10.1007/s10495-009-0446-3.
 1169 [117] A. Charnot, D. Gouveia, S. Aycirix, J. Lemoine, J. Armengaud, C. Almunia, A. Chaumot, O.
 1170 Geffard, A. Salvador, On-Line Solid Phase Extraction Liquid Chromatography-Mass
 1171 Spectrometry Method for Multiplexed Proteins Quantitation in an Ecotoxicology Test Specie:
 1172 *Gammarus fossarum*, *Journal of Applied Bioanalysis*. 4 (2018) 81–101.
 1173 doi:10.17145/jab.18.012.
 1174 [118] A. Charnot, D. Gouveia, J. Armengaud, C. Almunia, A. Chaumot, J. Lemoine, O. Geffard, A.
 1175 Salvador, Multiplexed assay for protein quantitation in the invertebrate *Gammarus fossarum*
 1176 by liquid chromatography coupled to tandem mass spectrometry, *Analytical and
 1177 Bioanalytical Chemistry*. 409 (2017) 3969–3991. doi:10.1007/s00216-017-0348-0.
 1178
 1179

1180 List of tables and figures

1181 **Figure 1** : Proteogenomic workflow for the identification of proteins from the hemolymph of
 1182 *Dreissena polymorpha*. (Picture of zebra mussel palleal cavity: I. Bonnard)

1183 **Figure 2.** Information regarding the protein identities. (A) Venn diagram of the number of proteins in
 1184 hemocyte and plasma as well overlapped proteins; (B) Venn diagram of the number of immune related
 1185 proteins in hemocyte and plasma as well overlapped proteins.
 1186

1187 **Figure 3.** Top Hit species distribution with NCBI nr Db, only species assigned to more than 8
 1188 sequences are represented

1189 **Figure 4.** Gene ontology analysis of the proteins mapped with the Go term “Immune system process”
 1190 in hemocyte (blue bars) and plasma (orange bars). (A) GO-terms related to biological process; (B)
 1191 GO-term related to Molecular function.
 1192

1193 **Table 1.** Abundance of Immune proteins identified in hemocytes and plasma compartments

1194

1195

1196 **Appendix A. Supplement Material**

1197 **Supplement Table S1. Identified proteins from the hemocytes.**

1198 Protein scores, percent peptide coverage of protein, spectral count of biological replicates,
1199 NSAF values, functional annotations and GO-term assignment are displayed for each protein.

1200

1201 **Supplement Table S2. Identified proteins from the plasma.**

1202 Protein scores, percent peptide coverage of protein, spectral count of biological replicates,
1203 NSAF values, functional annotations and GO-term assignment are displayed for each protein.

1204

1205 **Supplement Table S3. Conserved domain search of proteins from the hemolymph.**

1206 **Supplement Table S4. List of peptides identified from the hemocytes.**

1207 Peptide sequences, peptide scores and peptide modifications are displayed for each protein
1208 and each sample of hemocytes.

1209

1210 **Supplement Table S5. List of peptides identified from the plasma.**

1211 Peptide sequences, peptide scores and peptide modifications are displayed for each protein
1212 and each sample of plasma

1213

1214 **Conflicts of Interest**

1215 The authors declare no conflict of interest.

1216

1217 **Abbreviations**

1218 MAMP : Microbe-associated molecular pattern

1219 DAMP : Damage-associated molecular pattern

1220 RNA-seq : mRNA sequencing

1221 MS/MS: Tandem mass spectrometry

1222 NSAF : Normalized spectral abundance factor

1223 SC : Spectral count

1224 NCBI : National Center for Biotechnology Information

1225 GO : Gene Ontology

1226 PSM : Peptide spectrum matches

1227 Db : Data base

1228 PRR : Pattern recognition receptors

1229 CTLs : C-type lectins

1230 FREPs : Fibrinogen-related proteins

1231 FreD : Fibrinogen-related domain

1232 MPR : Mannose-6-phosphate receptor

1233 GBP : Beta-1,3-glucan-binding protein

1234 PGRP : Peptidoglycan-recognition protein

1235 AIF-1 : Allograft inflammatory factor 1

1236 MIF : Macrophage migration inhibitory factor

1237 MAPK : Mitogen-Activated Protein Kinase

1238 JNK : c-jun N-terminal kinase
1239 ERKs : Extracellular signal-regulated kinases
1240 IFN : Interferon
1241 STAT : Signal transducers and activators of transcription
1242 NF- κ B : Nuclear factor kappa B
1243 PKC : Protein kinase C -type
1244 PI3K : Phosphatidylinositol 4,5-bisphosphate 3-kinase
1245 Arp2/3 : Actin-related protein 2/3
1246 CDC42 : Cell Division Cycle 42
1247 Rac1 : Ras-related C3 botulinum toxin substrate
1248 NWASP : Neural Wiskott-Aldrich syndrome protein
1249 AP : Adaptor protein
1250 AMP : Antimicrobial peptide
1251 TLP : Thaumatin-like protein
1252 BPI : Bactericidal/permeability-increasing protein
1253 cyto-c : Cytochrome-c
1254 AIF : Apoptosis Inducing Factor
1255 IAP : Inhibitor of apoptosis

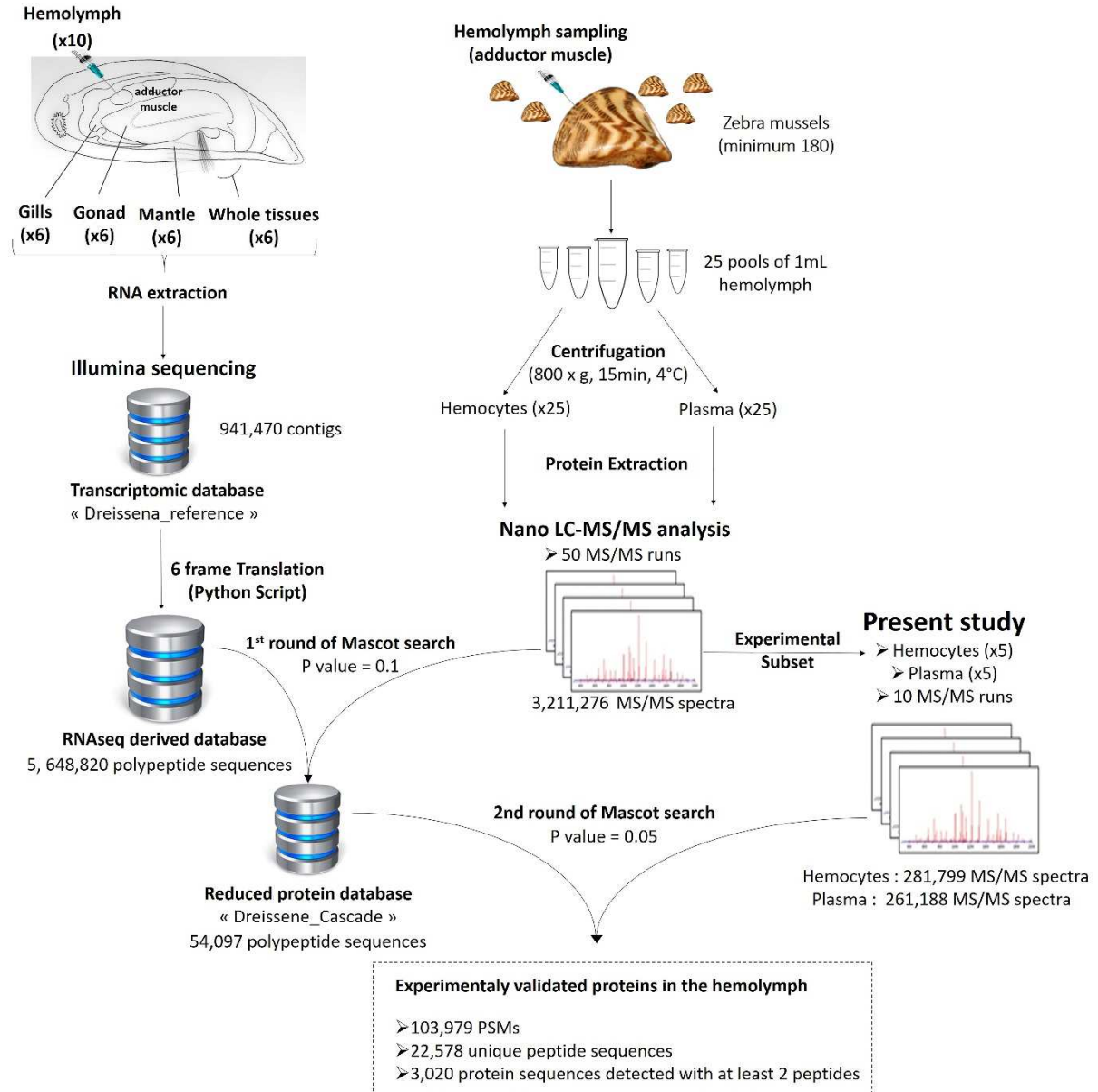


Figure 1. Proteogenomic workflow for the identification of proteins from the hemolymph of *Dreissena polymorpha*. (Picture of zebra mussel palleal cavity: I. Bonnard)

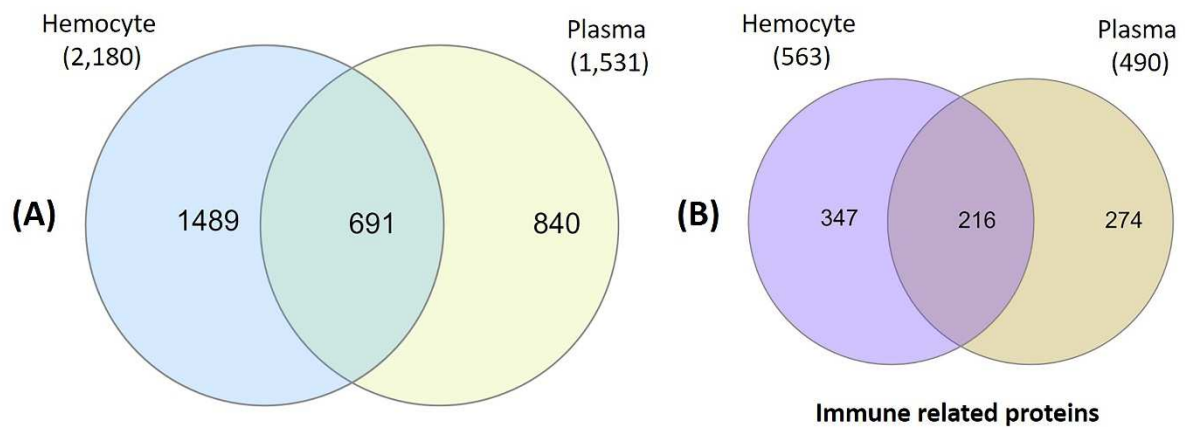


Figure 2. Information regarding the protein identities. (A) Venn diagram of the number of proteins in hemocyte and plasma as well overlapped proteins; (B) Venn diagram of the number of immune related proteins in hemocyte and plasma as well overlapped proteins.

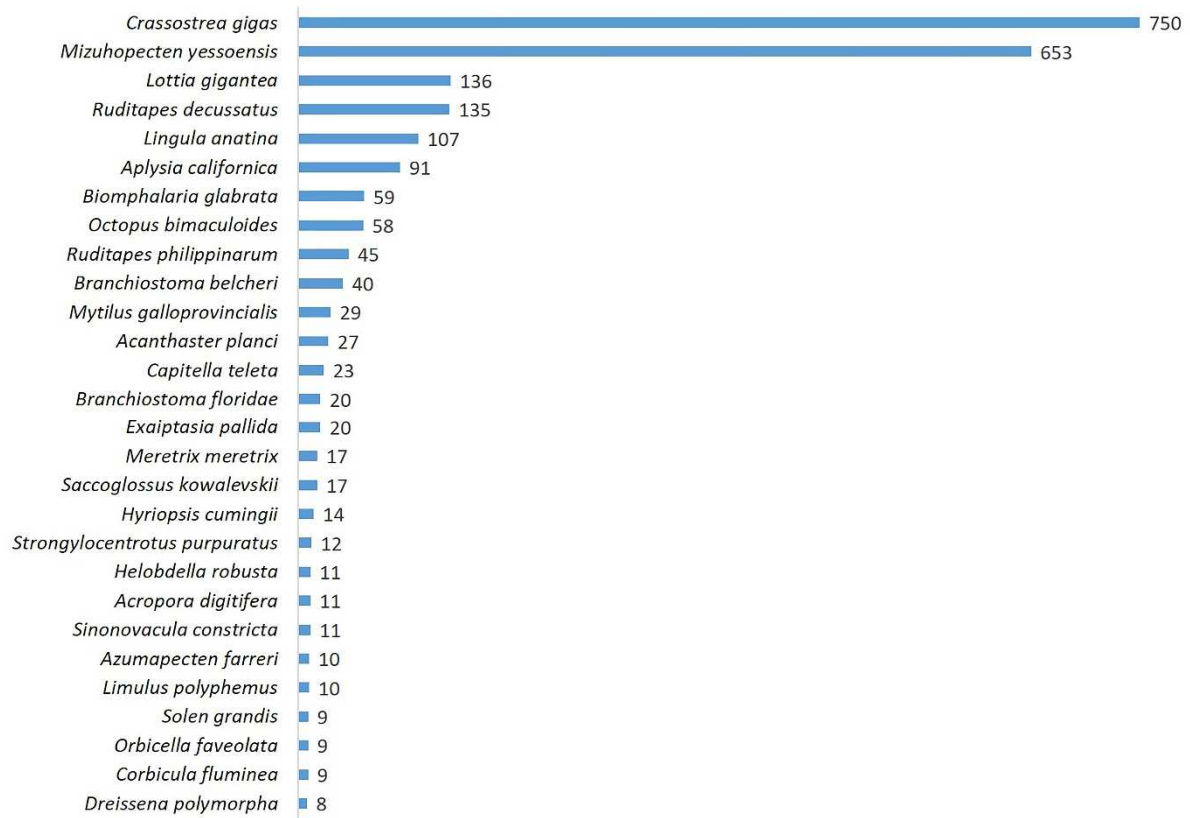


Figure 3. Top Hit species distribution with NCBI nr Db, only species assigned to more than 8 sequences are represented

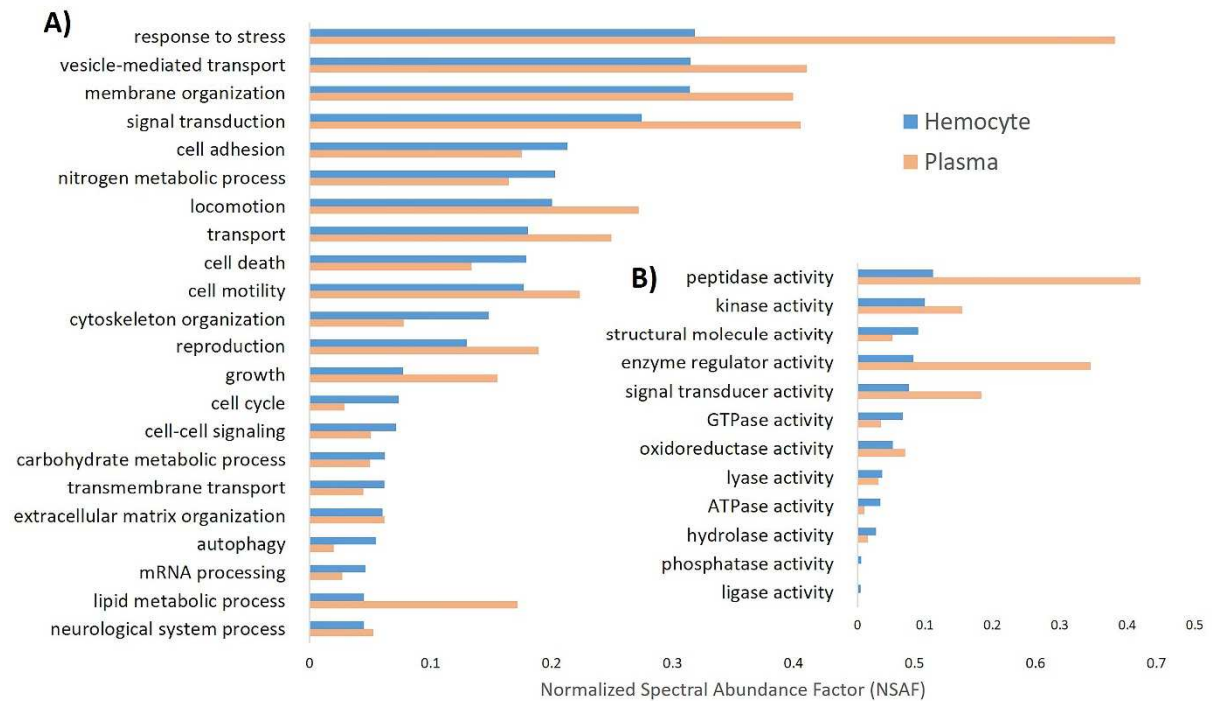
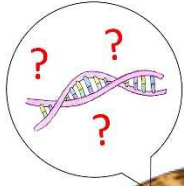


Figure 4. Gene ontology analysis of the proteins mapped with the Go term “Immune system process” in hemocyte (blue bars) and plasma (orange bars). (A) GO-terms related to biological process; (B) GO-terms related to Molecular function.

Table 1. Abundance of Immune proteins identified in hemocytes and plasma compartments

Protein functions	Hemocytes		Plasma	
	Spectral counts	% NSAF	Spectral counts	% NSAF
Recognition of micro-organisms				
Natterin-like protein	804	2.19	898	2.69
Galectins	588	1.96	484	1.41
MytiLec-like proteins	185	0.86	171	0.94
Fibrinogen-related proteins	222	0.32	428	0.74
Beta-1,3-glucan-binding proteins (BGBP)	269	0.17	203	0.13
C-type lectins	169	0.14	956	1.34
Peptidoglycan-recognition proteins	36	0.05	378	0.53
Cation-dependent mannose-6-phosphate receptors	-	-	57	0.01
Protein involved in immune signalling				
Mitogen-activated protein kinases	320	0.21	80	0.05
Proinflammatory cytokines	84	0.10	43	0.06
Interferon-Responsive Factors	97	0.06	-	-
PKC and PI3k related proteins	77	0.03	80	0.21
Nuclear factor kappa B pathway	37	0.01	-	-
Internalization of microorganisms				
GTPases of the Rho family	1890	1.16	325	0.24
Actin-related protein 2/3 complex	1017	0.77	309	0.23
Integrin related proteins	941	0.57	85	0.05
Clathrin-mediated endocytosis	349	0.10	37	0.01
Antimicrobial peptides				
Defensin	13	0.04	-	-
Thaumatins	201	0.23	149	0.18
lysozymes	7	0.01	51	0.08
Histones	1323	2.64	58	0.14
Proteases and Protease Inhibitors				
Cathepsins	455	1.32	110	0.29
Other proteases	363	0.19	1251	1.53
SERine Proteinase INhibitors (serpins)	31	0.04	-	-
Alpha-2-macroglobulins	756	0.47	4468	4.24
Cystatins	238	0.99	162	0.71
Complement system				
Complement component C3	585	0.96	9771	16.45
c1q domain containing proteins	194	0.68	1371	2.12
Complement component C2/factor B	72	0.17	1392	3.11
Other proteins related to complement system	12	0.05	310	1.16
Apoptosis				
Apoptosis regulators	460	0.27	86	0.06
Caspases	130	0.09	28	0.03

Non model species



Zebra mussel

Nano LC-MS/MS

Hemolymph proteome

Proteogenomic

RNA sequencing

