



HAL
open science

High polymorphism in big defensin gene expression reveals presence-absence gene variability (PAV) in the oyster *Crassostrea gigas*

Rafael D. Rosa, Pascal Alonso, Adrien Santini, Agnes Vergnes, Evelyne Bachere

► To cite this version:

Rafael D. Rosa, Pascal Alonso, Adrien Santini, Agnes Vergnes, Evelyne Bachere. High polymorphism in big defensin gene expression reveals presence-absence gene variability (PAV) in the oyster *Crassostrea gigas*. *Developmental and Comparative Immunology*, 2015, 49 (2), pp.231-238. 10.1016/j.dci.2014.12.002 . hal-04200548

HAL Id: hal-04200548

<https://hal.science/hal-04200548>

Submitted on 12 Oct 2023

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 4.0 International License

High polymorphism in big defensin gene expression reveals presence-absence gene variability (PAV) in the oyster *Crassostrea gigas*

Rosa Rafael D.^{1,*}, Alonso Pascal¹, Santini Adrien¹, Vergnes Agnes¹, Bachere Evelyne¹

¹ Ifremer, CNRS, Université Montpellier 2, IRD, and Université Montpellier 1 (UMR5119 - Ecology of Coastal Marine Systems), Place Eugène Bataillon, CC80, 34095 Montpellier, France

* Corresponding author : Rafael D. Rosa, tel.: +55 48 3721 6163 ; Fax: +55 48 3721 5528 ; email address : rafael.d.rosa@ufsc.br

Abstract :

We report here the first evidence in an invertebrate, the oyster *Crassostrea gigas*, of a phenomenon of Presence-Absence Variation (PAV) affecting immune-related genes. We previously evidenced an extraordinary interindividual variability in the basal mRNA abundances of oyster immune genes including those coding for a family of antimicrobial peptides, the big defensins (Cg-BigDef). Cg-BigDef is a diverse family composed of three members: Cg-BigDef1 to -3. Here, we show that besides a high polymorphism in Cg-BigDef mRNA expression, not all individual oysters express simultaneously the three Cg-BigDefs. Moreover, in numerous individuals, no expression of Cg-BigDefs could be detected. Further investigation at the genomic level revealed that in individuals in which the transcription of one or all Cg-BigDefs was absent the corresponding Cg-bigdef gene was missing. In our experiments, no correlation was found between Cg-bigdef PAV and oyster capacity to survive *Vibrio* infections. The discovery of P-A immune genes in oysters leads to reconsider the role that plays the immune system in the individual adaptation to survive environmental, biotic and abiotic stresses.

Highlights

► Cg-BigDefs display a high interindividual polymorphism in basal mRNA expression. ► All three Cg-BigDefs are not simultaneously expressed in a single oyster. ► Lack of Cg-BigDef mRNA expression is likely due to the absence of the corresponding gene. ► First report in an invertebrate of Presence-Absence Variation affecting immune genes.

Keywords : mollusk bivalve, antimicrobial peptide, polymorphism, gene copy number variation, high-throughput microfluidic RT-qPCR, *Vibrio* infection

46 1. Introduction

47 Understanding the mechanisms of resistance or susceptibility to infections in the oyster
48 *Crassostrea gigas* should contribute to identify the causes of mass mortalities that affect this
49 marine mollusk species worldwide. In this context of multifactorial disease, significant
50 progresses were made in characterizing the oyster immune system and responses to biotic and
51 abiotic environment factors. Particular attention has been paid to the mechanisms involved in
52 the capability of the oyster to overcome infections (de Lorgeril et al., 2011; Schmitt et al.,
53 2012) but also to the genetic bases that govern oyster survival (Rosa et al., 2012).

54 High-throughput transcriptomic studies have revealed in *C. gigas* great individual
55 polymorphism in the expression of many immune-related genes (Rosa et al., 2012). This
56 supported previous evidence that *C. gigas* is a highly polymorphic and intron-rich species in
57 terms of number of SNPs and repetitive DNA sequences (Kim et al., 2011; Saavedra and
58 Bachère, 2006; Sauvage et al., 2007). Thus, besides sequence polymorphism, the oyster
59 displays great polymorphism of expression that can be related to gene Copy Number
60 Variation (CNV) in genome contents between different individuals (Rosa et al., 2012; Schmitt
61 et al., 2013). CNV are known to be of fundamental importance for genetic variation in
62 populations, evolution of genome complexity and evolution of genes with novel functions.
63 Many studies have looked for associations between the copy number of genes and disease
64 susceptibility (reviewed in Cantsilieris and White, 2013). For instance, in human, whereas
65 extensive copy number of β -defensin antimicrobial peptide has been associated to the risk of
66 psoriasis (Hollox et al., 2008), low copy number of β -defensin has been reported to
67 predispose to Crohn disease (Fellermann et al., 2006).

68 Surprisingly, our high-throughput transcriptomic studies have also revealed numerous
69 genes whose basal levels of expression could not be detected in some individuals (Rosa et al.,
70 2012). Those genes were shown to be involved in immune responses such as genes encoding

71 members of the big defensin antimicrobial peptide family, *Cg-BigDef* (Rosa et al., 2011).
72 Oyster big defensins (*Cg-BigDef*) have been shown to be a diverse antimicrobial peptide
73 family composed of three members (*Cg-BigDef1*, *Cg-BigDef2* and *Cg-BigDef3*) that are
74 encoded by distinct genomic sequences (Rosa et al., 2011). An interesting particularity of the
75 oyster *Cg-BigDef* family is that their members follow different patterns of gene regulation
76 upon microbial challenge. Indeed, whereas *Cg-BigDef1* and *Cg-BigDef2* are up-regulated in
77 circulating hemocytes in response to pathogenic *Vibrio* infection, *Cg-BigDef3* is constitutively
78 expressed and non-regulated (Rosa et al., 2011). In this study, an unusual variability in
79 mRNA expression was evidenced in oyster sampling for all *Cg-BigDefs*. As *Cg-BigDef1* and -
80 2 transcriptions are highly induced by microbial molecular patterns (Rosa et al., 2011), it was
81 reasonable to expect that these molecules could be implicated in preventing from
82 development infectious processes. This prompted us to further analyze such polymorphism of
83 expression in *C. gigas*.

84 Here, a detailed analysis of *Cg-BigDef* mRNA expression at individual level confirmed
85 firstly a high interindividual variability in basal gene expression for the three *Cg-BigDef*
86 forms. In addition, we found that the three *Cg-BigDefs* are not simultaneously transcribed in a
87 single oyster. More surprisingly, in some individual oysters, basal mRNA levels were not
88 detected for any *Cg-BigDefs*. Besides, we evidenced that the absence of *Cg-BigDef*
89 transcription is likely associated to the lack of *Cg-bigdef* genes, revealing the extreme form of
90 CNV named Presence-Absence Variation (PAV). Finally, the Presence-Absence Variation of
91 *Cg-bigdefs* in the oyster genome was shown to be not related to a better capacity of the
92 oysters to survive experimental *Vibrio* infections.

93

94 **2. Materials and Methods**

95 **2.1. Oysters and experimental *Vibrio* infections**

96 Adult two-year old *Crassostrea gigas* oysters were purchased from a local oyster farm
97 in Mèze (Gulf of Lion, France) and kept in aquaria containing filtered seawater at 15°C at the
98 Aquaculture Experimental Platform of Ifremer, Palavas-Les-Flots. During two weeks, animals
99 were fed twice with a mixture of live microalgae (*Nannochloropsis oculata*: 2.5×10^5 cells/ml;
100 *Tetraselmis suecica*: 2.5×10^2 cells/ml). Following the acclimatation period, 90 oysters were
101 individually tagged and hemolymph (~500 µl per animal) was withdrawn from the adductor
102 muscle without causing the death of the animals. Hemocyte samples were individually
103 collected by centrifugation (1,500 ×g, 15 min, 4°C), homogenized in 1 ml of TRIzol[®] reagent
104 (Invitrogen) and frozen at -20°C until RNA extraction. After hemolymph collection, animals
105 were placed in glass tanks (20 animals per tank) and allowed to recover for 8 days prior to
106 experimental infections. Then, animals were intramuscularly injected with 5×10^8 CFU/animal
107 of the virulent *V. tasmaniensis* LGP32 (= *V. splendidus* LGP32). Mortalities were monitored
108 daily and individually tagged animals that did not survive were noted and discarded. After the
109 end of acute mortalities (96 h post-infection), stored hemolymph samples were categorized
110 according to oyster survival (surviving *versus* non-surviving). The standardization of the
111 experimental infections and the preparation of the bacterial inoculum were performed as
112 previously described (Duperthuy et al., 2010). All experimentations were performed
113 according to the Ifremer animal care guideline and policy.

114 Hemocyte samples from 73 additional oysters (naive animals) were obtained as
115 described above, homogenized in 1 ml of TRIzol[®] reagent (Invitrogen) and frozen at -20°C
116 until RNA extraction. From those, 23 oysters were randomly chosen, removed from their
117 shells and immediately frozen in liquid nitrogen. Frozen whole-body oysters were
118 individually ground to fine powder for further extractions of total RNA and genomic DNA in
119 same individuals.

120

121 2.2. Genomic DNA isolation

122 Genomic DNA (gDNA) was isolated individually as previously described (Schmitt et
123 al., 2010). Briefly, 50-100 mg of oyster powder were incubated overnight at 55°C in 500 µl of
124 DNA extraction buffer (100 mM NaCl, 10 mM Tris-HCl pH 8, 25 mM EDTA pH 8, 0.5%
125 SDS, 0.1 mg/ml protease K). Following a step of phenol/chloroform extraction, gDNA
126 samples were precipitated by addition of cold ethanol and treated with 50 µg/ml RNase A
127 (Invitrogen) for 30 min at 37°C to eliminate contaminating RNA. Quantification and quality
128 of gDNA samples were assessed by spectrophotometry (NanoDrop ND-1000 Thermo
129 Scientific) and 0.8% agarose gel electrophoresis, respectively.

130

131 2.3. Total RNA extraction and cDNA synthesis

132 Total RNA was extracted individually by using TRIzol[®] reagent (Invitrogen) according
133 to the manufacturer's instructions. RNA samples were then treated with DNase I (Invitrogen)
134 for 15 min at room temperature to eliminate contaminating genomic DNA. After DNase I
135 inactivation (10 min at 65°C), samples were precipitated with 0.3 M sodium acetate. RNA
136 amount and quality were assessed by spectrophotometric analysis and the integrity of total
137 RNA was analyzed by 0.8% agarose gel electrophoresis.

138 Following heat denaturation (70°C for 5 min), reverse transcription was performed
139 using 250 ng of purified total RNA with 50 ng/µl oligo(dT)₁₂₋₁₈ in a 20-µl reaction volume
140 containing 1 mM dNTPs, 1 unit of RNaseOUT Ribonuclease and 200 units M-MLV reverse
141 transcriptase in reverse transcriptase buffer according to the manufacturer's instructions
142 (Invitrogen). Attempting to increase the number of cDNA templates, a supplementary step of
143 preamplification was performed using 1.25 µl of cDNA, 1.25 µl of a mixture of forward and
144 reverse primers (200 nM each) and 2.5 µl of 2× TaqMan[®]PreAmpMaster Mix (Applied
145 Biosystems). The specific primer pairs for each Cg-BigDef form are listed in **Table 1**. PCR

146 cycling conditions were as follows: polymerase activation at 95°C for 10 min, followed by 14
147 cycles of 95°C for 15 s and 60°C for 4 min. Pre-amplification products were diluted to 1:5 in
148 TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5) and stored at -20°C.

149

150 **2.4. High-throughput RT-qPCR using 96.96 microfluidic dynamic arrays**

151 The relative gene expression was measured by using the high-throughput microfluidic
152 RT-qPCR platform BioMark™ (Fluidigm) (Jang et al., 2011). The sample reaction mixtures
153 were performed in a final volume of 5 µl containing 1.25 µl of pre-amplified cDNA (diluted
154 1:5), 2.5 µl of 2× TaqMan® Gene Expression Master Mix (Applied Biosystems), 0.25 µl of
155 20× DNA Binding Dye Sample Loading Reagent (Fluidigm), 0.25 µl of 20× EvaGreen
156 (Biotium) and 0.75 µl of TE buffer. Primer reaction mixtures were made in the same volume
157 of 5 µl containing 2.5 µl of 2× Assay Loading Reagent (Fluidigm), 1.25 µl of 20 µM of
158 forward and reverse primer mix and 1.25 µl of TE buffer. Both sample and primer reaction
159 mixtures were loaded into the dynamic array chip that was subsequently placed on the
160 NanoFlex™ 4-IFC Controller for loading and mixing. After approximately 50 min, the chip
161 was transferred to the BioMark™ Real-Time PCR System.

162 The cycling program used consisted of 10 min at 95°C followed by 40 cycles of 95°C
163 for 15 s and 1 min at 60°C. Melting curves analysis was performed after completed RT-qPCR
164 collecting fluorescence between 60-95°C at 0.5°C increments. Data were analyzed using the
165 BioMark™ Real-time PCR analysis software to obtain Cq values. Results are presented as
166 changes in relative expression normalized with the Cq geometric mean value of the eukaryotic
167 translation elongation factor 1-alpha (*Cg-ef1a*: AB122066), the ribosomal protein L40 (*Cg-*
168 *rpl40*: FP004478) and the 40S ribosomal protein S6 (*Cg-rps6*: CU686508) as endogenous
169 reference genes (Livak and Schmittgen, 2001). Statistical significance between expression

170 detection of *Cg-BigDefs* and oyster capacity to survive *V. tasmaniensis* LGP32 infections was
171 determined by ANOVA test at $p < 0.05$.

172

173 **2.5. Genotyping of *Cg-bigdefs***

174 PCR amplifications for *Cg-bigdef* genotyping were conducted in a 25- μ l reaction
175 volume using 1 μ l of synthesized cDNA or 30 ng of purified gDNA as template. Primer
176 sequences are listed in **Table 1**. PCR conditions were as follows: 10 min at 94°C, then 30
177 cycles at 94°C for 1 min, 56°C for 1 min, 72°C for 3 min and a final elongation step of 72°C
178 for 10 min. PCR products were analyzed by electrophoresis on 1.5% agarose gel stained with
179 ethidium bromide. The amplification products were cloned into a pCR[®]-Blunt II-TOPO[®]
180 cloning vector using a Zero Blunt[®] TOPO[®] PCR cloning kit (Invitrogen). The positive
181 recombinant clones were identified by colony PCR and were sequenced in both directions.

182

183 **2.6. Sequence data analysis**

184 Obtained sequences were analyzed for similarities using BLASTX at the National
185 Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov>). The multiple
186 alignments were generated by using the ClustalW2 Multiple Alignment Program (ClustalW2;
187 <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The phylogenetic analyses based on the full-
188 length cDNA and genomic sequences of *C. gigas* big defensins were performed using the
189 Neighbour-Joining method with the software MEGA version 4 (Tamura et al., 2007).
190 Bootstrap sampling was reiterated 1,000 times.

191

192 **3. Results**

193 **3.1. Oyster big defensins display a high interindividual polymorphism in basal gene**
194 **expression**

195 The basal expression levels of the three big defensins were first analyzed by a high-
196 throughput RT-qPCR device in 163 individual oysters (**Fig. 1; Appendix A**). Specific primers
197 were used for each *Cg-BigDef* form. The RT-qPCR analyses evidenced a high interindividual
198 variability for the three *Cg-BigDef* transcripts. The *Cg-bigdef1* and *Cg-bigdef2* forms whose
199 expression is inducible showed lower basal expression compared to the *Cg-bigdef3* form,
200 constitutively expressed. Figure 2A shows the relative mRNA abundance of *Cg-BigDefs* in
201 ten individual oysters which are representatives of individual differences in basal gene
202 expression of the 163 oysters analyzed in this study. Depending on individuals, differences
203 reached up to 30, 27 and 7-fold for *Cg-BigDef1*, *Cg-BigDef2* and *Cg-BigDef3*, respectively
204 (**Fig. 1; Fig. 2A**). Besides, the three *Cg-BigDef* forms were never found to be simultaneously
205 expressed among the 163 oysters analyzed (**Table 2**). Interestingly, only one or two forms of
206 *Cg-BigDefs* were shown to be expressed in a same individual. *Cg-BigDef1* and *Cg-BigDef2*
207 were seen in 81 and 47 individuals, respectively, while *Cg-BigDef3* was expressed in 21
208 oysters only. Similar results were obtained when using different specific primer pairs for each
209 *Cg-BigDef* form (**Table 1**).

210 More surprisingly, among the 163 oysters, 39 did not express any *Cg-BigDefs* (**Table**
211 **2**). This lack of expression was further confirmed by RT-qPCR analyses using two universal
212 primer pairs common to all *Cg-BigDef* sequences. As a control to these RT-qPCR analyses,
213 the expression of three reference genes (*Cg-rpl40*, *Cg-ef1a* and *Cg-rps6*) was found to be
214 uniform across all tested individuals (**Appendix A**). These results evidenced a high
215 interindividual variability in gene expression of *C. gigas* big defensins, including cases where
216 their basal expression is fully absent in some individuals.

217

218 **3.2. Lack of *Cg-BigDef* expression is likely associated to the absence of encoding genes in**
219 **the oyster genome**

220 In light of the degree of polymorphism in gene expression observed for *Cg-BigDefs*, we
221 further investigated the relationships between the expression and presence of the three *Cg-*
222 *bigdef* genes in 23 individuals from which we have extracted both mRNA and gDNA.
223 Detection of transcripts was performed by RT-qPCR and detection of corresponding genes by
224 conventional PCR with specific primer pairs for each form (**Table 1**). The expression profiles
225 obtained for these 23 individuals confirmed the high variability in basal expression levels
226 observed above by high-throughput RT-qPCR analyses (**Fig. 2A**; **Appendix A**). Figure 2B
227 shows the amplification of *Cg-bigdef* genes in ten individual oysters which are representatives
228 of the 23 oysters analyzed. Interestingly, genes encoding a given *Cg-BigDef* were only
229 amplified in individuals expressing the corresponding form at basal levels (**Fig. 2B**). The
230 accuracy of these results was further confirmed by using an additional primer pair specific for
231 the different *Cg-bigdefs* and two universal primer pairs (common to all forms). Besides, as a
232 control of the integrity of extracted gDNA samples, *Cg-ef1a* gene was successfully amplified
233 in all oysters analyzed (**Fig. 2B**).

234

235 **3.3. *Cg-BigDef* forms are likely encoded by three distinct genes**

236 We performed an exhaustive gDNA and cDNA cloning and sequencing of the *Cg-*
237 *bigdefs* from four individuals expressing one or two forms, first to ensure the authenticity of
238 the PCR amplifications but also to gain insights into the genomic organization of the big
239 defensin family. We sampled one individual that transcribed simultaneously *Cg-BigDef1* and
240 -2 as well as one individual that was shown to transcribe both *Cg-BigDef1* and -3. Two
241 additional individuals expressing only *Cg-BigDef2* or -3 were also sampled. Both gene and
242 cDNA sequences were obtained from these four individuals where the *Cg-BigDefs* were
243 detected at the level of transcripts and genes. The obtained sequences were deposited in
244 GenBank under the accession numbers JN251121 to JN251132.

245 Multiple alignments of the *Cg-BigDef* gDNA and cDNA sequences obtained in single
246 individuals confirmed that each oyster big defensin form is transcribed by a separated gene.
247 Indeed, whereas *Cg-bigdef1* and *Cg-bigdef2* genes included two exons interrupted by a single
248 intron, an additional intron and exon were observed in *Cg-bigdef3* upstream the first exon
249 common to the other *Cg-bigdef* genes (**Fig. 3**). Interestingly, in one individual in which both
250 *Cg-BigDef1* and -2 were expressed, two different *Cg-bigdef1* sequences were identified in
251 addition to the *Cg-bigdef2* gene (**Fig. 4**). In this oyster, the *Cg-bigdef1* genes differed in the
252 length of the intron sequence and by the presence of two nucleotide substitutions (**Fig. 4**).
253 Moreover, for the other individuals analyzed, single nucleotide polymorphisms (SNP) were
254 found between transcripts and genes, which was evocative of two different alleles of a same
255 *Cg-bigdef* gene. Thus, each *Cg-bigdef* would correspond rather to a distinct gene than to
256 alleles from a same *Cg-bigdef* locus.

257 Sequence analysis revealed that, for each *Cg-bigdef* gene, the length of exons was well
258 conserved among individuals whereas the introns lengths were variable. According to
259 individuals, SNP were identified in both intron and exon sequences for a same *Cg-bigdef*
260 gene. Additionally, introns from all *Cg-bigdef* genes included some microsatellite sequences
261 (such as CTAT, CT and CA) and the canonical GT/AG splice site junctions (**Fig. 3**). The
262 intron sequences of *Cg-bigdef3* appeared to differ considerably in both length and nucleotide
263 composition from those of *Cg-bigdef1* and *Cg-bigdef2* genes (**Fig. 3**).

264 A phylogenetic tree was built with the nucleotide sequence of six *Cg-bigdef* gene
265 sequences obtained here and with *Cg-bigdef* gene sequences previously deposited in GenBank
266 (Rosa et al., 2011). In this tree, *Cg-bigdef1* and *Cg-bigdef2* genes clustered in a same clade
267 distinct from the *Cg-bigdef3* gene (**Fig. 5**). An additional phylogenetic tree constructed from
268 cDNA sequence data confirmed that *Cg-BigDef1* and *Cg-BigDef2* forms are more related to
269 each other than to *Cg-BigDef3* (**Fig. 5**). Altogether, results from sequencing and phylogenetic

270 analyses strongly suggest that *Cg-BigDefs* constitute a multigenic family of antimicrobial
271 peptides in *C. gigas* oysters.

272

273 **3.4. Presence-absence variation (PAV) in *Cg-BigDef* gene expression is not associated to** 274 **oyster capacity to survive virulent *Vibrio* infections**

275 In order to investigate whether the presence or absence of *Cg-BigDefs* was associated to
276 an inherent capacity of oysters to survive infections, we have quantified the mRNA levels of
277 all three *Cg-BigDefs* (*Cg-BigDef1*, *Cg-BigDef2* and *Cg-BigDef3*) in 90 individual animals
278 before a challenge with an oyster pathogen. By using a nonlethal sampling method prior to the
279 experimental infections (Rosa et al., 2012), we have assessed the basal transcription profile of
280 all three *Cg-BigDefs* in oysters able (S) and not able (NS) to survive infections by the virulent
281 *V. tasmaniensis* LGP32 (= *V. splendidus* LGP32). Three independent experimental infections
282 were conducted and hemolymph samples from 45 oysters of each survival capacity phenotype
283 were obtained.

284 The detection of *Cg-BigDef* transcripts and their mRNA abundance were compared in
285 individual oysters from both S and NS phenotypes by high-throughput microfluidic RT-qPCR
286 analyses. Transcripts of *Cg-BigDef1* were detected in 25 individuals from both S and NS
287 phenotypes whereas transcripts of *Cg-BigDef2* were detected in 10 and 12 individuals from S
288 and NS phenotypes, respectively (**Table 3**). *Cg-BigDef3* transcripts were detected in 4
289 individuals from the S phenotype and in 2 from the NS phenotype. For the S phenotype, 27
290 and 6 individuals expressed, respectively, only one or two *Cg-BigDefs* (*Cg-BigDef1* and *Cg-*
291 *BigDef2* or *Cg-BigDef1* and *Cg-BigDef3*) and 12 oysters did not express any *Cg-BigDefs*
292 (**Table 3**). For the NS phenotype, 4 and 31 oysters expressed, respectively, only one or two
293 *Cg-BigDefs* (*Cg-BigDef1* and *Cg-BigDef2* or *Cg-BigDef1* and *Cg-BigDef3*) and 10 oysters did
294 not express any *Cg-BigDefs*. No oysters expressing simultaneously *Cg-BigDef2* and *Cg-*

295 *BigDef3* or all three *Cg-BigDefs* were found in both S and NS phenotypes (**Table 3**). No
296 significant differences were found in the detection of *Cg-BigDefs* according to the S and NS
297 phenotypes, namely according to the oyster capability to survive virulent *Vibrio* infections
298 ($p>0.05$). Additionally, no significant differences in basal gene expression levels of the three
299 *Cg-BigDefs* were found between individuals from both survival capacity phenotypes.

300

301 **4. Discussion**

302 Taking advantages of the development of high-throughput genomic technologies, we
303 have explored the genetic determinants for *Crassostrea gigas* capacity to survive infections at
304 the individual level. These analyses highlighted an extraordinary interindividual
305 polymorphism in the basal expression of genes in this species (Rosa et al., 2012). Indeed,
306 most of the transcriptome profilings are performed with oyster sample pools, which hides
307 great information about oyster genomic diversity and complexity. Here, high-throughput RT-
308 qPCR analyses at the individual level revealed a high variability in basal gene expression
309 regulation of antimicrobial peptides belonging to *Cg-BigDef* family. Basal mRNA amounts of
310 both the inducibly expressed *Cg-BigDef1* and *Cg-BigDef2* and the constitutively-expressed
311 *Cg-BigDef3* forms appeared to be highly variable among individual oysters. For instance,
312 variations up to 30-fold were found in *Cg-BigDef1* expression according to individuals. This
313 interindividual polymorphism in *Cg-BigDef* transcription levels may reflect genetic structure
314 variation in gene copy numbers (CNV) as recently reported for other oyster AMPs, *Cg-Defs*
315 and *Cg-Prps* (Schmitt et al., 2013).

316 When analyzing the heterogeneity of transcription profiles of the big defensins, we
317 showed that no expression was detected in some oysters and that the lack of basal expression
318 was likely associated to the absence of the corresponding encoding sequences at the genome
319 level. Such results highlight structural variations in oyster genomes and particularly gene

320 Presence-Absence variation (PAV), a phenomenon described as the extreme case of Copy
321 Number Variation (CNV) in which a large DNA segment present in some individuals can be
322 entirely missing in others. To our knowledge, this is the first evidence of PAV in oysters.

323 Considering that *C. gigas* oyster displays a high degree of DNA sequence
324 polymorphism and heterozygosity (Sauvage et al., 2007; Schmitt et al., 2010; Taris et al.,
325 2008), a rigorous accuracy has been required in the choice of the primers used in this study.
326 Two primer pairs specific for each of the three *Cg-BigDefs* have been designed as well as two
327 primer pairs common for all three forms. The relationship between the presence-absence (P-
328 A) of transcripts and the presence-absence of corresponding encoding gene has been
329 confirmed in 23 individuals by different PCR analyses using eight different primer pairs
330 spread across the entire sequences. Besides, the correspondence between the different genes
331 and their corresponding transcripts amplified has been assessed by nucleotide sequencing.
332 Thus, the non-amplification of any *Cg-BigDef* is likely due to the lack of genes instead of the
333 existence of null alleles commonly reported in oysters (Reece et al., 2004).

334 Here, we have improved also knowledge about oyster big defensin diversity by the
335 identification of novel variants for each *Cg-BigDef* form. Interestingly, in single individuals,
336 we evidenced the presence of different variants for a same *Cg-bigdef* gene, suggesting the
337 presence of different isoforms of a given *Cg-BigDef* in a same oyster. Likewise, the human
338 *DEF* locus has been shown to be highly variable among individuals (Taudien et al., 2004).
339 Variations in the human *DEF* locus include the presence of different alleles for a same *DEF*
340 gene, and both CNV and PAV phenomena.

341 PAV appeared to concern all *Cg-bigdefs* but more particularly the constitutively
342 expressed *Cg-bigdef3* gene that was recorded in less than 13% of the analyzed oyster
343 population. Importantly, since we sampled animals from a same oyster hatchery, we cannot
344 presume that the obtained frequency distribution of *Cg-bigdef* genes is representative of

345 natural oyster populations or representative of the species. The three *Cg-BigDefs* were never
346 seen simultaneously expressed in a single oyster, probably due to the relatively limited oyster
347 sampling size. Indeed, based on the occurrence frequency of the three *Cg-bigdefs* in this
348 oyster population, the probability of one individual carrying the three loci was calculated to be
349 less than 2%. Any of the three big defensin forms were detected in 23.9% (39/163) of the
350 individuals examined. A single oyster can express only one or up two *Cg-BigDefs*
351 concomitantly. Conversely, in *Litopenaeus* shrimp species all three members of the penaeidin
352 family (*Litvan* PEN2, *Litvan* PEN3 and *Litvan* PEN4) and all four members of the anti-
353 lipopolysaccharide factor (ALF) family (*Litvan* ALF-A, *Litvan* ALF-B, *Litvan* ALF-C and
354 *Litvan* ALF-D) are simultaneously transcribed in a single individual (Cuthbertson et al., 2002;
355 O’Leary and Gross, 2006; Rosa et al., 2013). Each penaeidin member is encoded by a
356 separated gene and the interindividual molecular diversity is mainly generated by DNA
357 polymorphism within each *PEN* gene locus (O’Leary and Gross, 2006). The simultaneous
358 expression of different members of a given antimicrobial peptide family in a single animal has
359 been shown to improve its antimicrobial defenses (Cuthbertson et al., 2002; Rosa et al., 2013).

360 Analyses of the *Cg-bigdef* gene sequences obtained from a single individual suggest that
361 each *Cg-BigDef* form is encoded by a separate gene and not by different alleles from a same
362 *Cg-bigdef* locus. Interestingly, the intron sequences of *Cg-bigdef3* were shown to be different
363 from those of *Cg-bigdef1* and *Cg-bigdef2* genes, suggesting an ancient divergence between
364 the constitutively and the inducibly expressed forms. Likewise to *Cg-BigDefs*, *Cg-Defs* and
365 *Cg-Prps* are also multigenic families whose the different members are encoded by separate
366 genes displaying distinct gene structures (Schmitt et al., 2010). Thus far, no transcription
367 profiling of *Cg-Defs* and *Cg-PRPs* have been performed at the individual level but it can be
368 assumed that PAV can also affect these effectors in oysters.

369 To date, intraspecific variations in gene contents such as PAV have never been reported
370 in oysters and any other non-model invertebrate species, probably because few genome
371 sequences are available. Indeed, whole *C. gigas* genome has been recently sequenced from
372 one individual (Zhang et al., 2012). It is noteworthy that we failed to evidence in the
373 published database any full sequences of big defensin encoding genes (Zhang et al., 2012).
374 The PAV phenomenon for big defensins or any other genes in *C. gigas* deserves to be fully
375 characterized by genetic mapping or through population-scale genome sequencing.

376 PAV is essentially known in human and particularly in plants in which this genetic
377 diversity is exploited in domestication programs. In human, PAV has been described for the
378 *DEFA3* gene encoding human neutrophil AMP HNP-3 (Linzmeier and Ganz, 2005).
379 Interestingly, the occurrence of *DEFA3* gene is variable depending on the human
380 geographical populations, and it has been suspected to be implicated with differences in
381 infectious disease susceptibility (Ballana et al., 2007). In our study, no correlation has been
382 shown between the P-A prevalence of the big defensins and the intrinsic capacity of the
383 oysters to survive experimental *Vibrio* infections. However, one can expect that P-A genes
384 others than those genes encoding AMP could have substantial involvement on oyster fitness
385 and health in the context of multifactorial disease that currently affects this species
386 worldwide.

387 In plants, regions of the genome encompassing PAV have been shown to contain gene
388 clusters involved in disease resistance (Zhang et al., 2014). Thus, it is assumed that PAV may
389 play important role in generating new pathogen resistances and they are now considered as
390 polymorphic markers for genetic studies and breeding (Wang et al., 2014; Zhang et al., 2014).
391 Overall, the great number of P-A genes observed in plants is recognized to play important
392 roles in individual adaptation to external pressure and environmental stresses (Haun et al.,
393 2011; Springer et al., 2009; Swanson-Wagner et al., 2010). Indeed, such structural variations

394 imply that the missing genes in some individuals are not necessary for survival but rather
395 contribute to the genome evolution and to the interplay with the environment (Marroni et al.,
396 2014). Undoubtedly, the first discovery of PAV that affect immune-related genes in oyster
397 open questions about the functional consequences of such genome structural variation on the
398 oyster immune responses but also on the individual capability to survive environmental, biotic
399 and abiotic stresses.

400

401 **Acknowledgments**

402 We are in debt with Frédérique Cerqueira, Jean-Luc Rolland and Marc Leroy for their
403 excellent technical assistance. We are also grateful to François Bonhomme for helpful
404 comments and Delphine Destoumieux-Garzón for critical reading of the manuscript. This
405 study received financial support from Ifremer and CNRS. RD Rosa was supported by a
406 doctoral founding from CNPq-Brazil. Data used in this work were partly produced through
407 molecular genetic analysis technical facilities of the SFR “Montpellier Environnement
408 Biodiversité” and the “Platform qPHD UM2/Montpellier GenomiX”.

409

410 **References**

- 411 Ballana, E., González, J.R., Bosch, N., Estivill, X., 2007. Inter-population variability of
412 DEFA3 gene absence: correlation with haplotype structure and population variability.
413 BMC Genomics 8, 14.
- 414 Cantsilieris, S., White, S.J., 2013. Correlating multiallelic copy number polymorphisms with
415 disease susceptibility. Hum. Mutat. 34, 1–13.
- 416 Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S., 2002. Diversity of the
417 penaeidin antimicrobial peptides in two shrimp species. Immunogenetics 54, 442–5.

- 418 de Lorgeril, J., Zenagui, R., Rosa, R.D., Piquemal, D., Bachère, E., 2011. Whole
419 transcriptome profiling of successful immune response to *Vibrio* infections in the oyster
420 *Crassostrea gigas* by digital gene expression analysis. PLoS One 6, e23142.
- 421 Duperthuy, M., Binesse, J., Le Roux, F., Romestand, B., Caro, A., Got, P., Givaudan, A.,
422 Mazel, D., Bachère, E., Destoumieux-Garzón, D., 2010. The major outer membrane
423 protein OmpU of *Vibrio splendidus* contributes to host antimicrobial peptide resistance
424 and is required for virulence in the oyster *Crassostrea gigas*. Environ. Microbiol. 12,
425 951–63.
- 426 Fellermann, K., Stange, D.E., Schaeffeler, E., Schmalzl, H., Wehkamp, J., Bevins, C.L.,
427 Reinisch, W., Teml, A., Schwab, M., Lichter, P., Radlwimmer, B., Stange, E.F., 2006. A
428 chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy
429 number predisposes to Crohn disease of the colon. Am. J. Hum. Genet. 79, 439-48.
- 430 Haun, W.J., Hyten, D.L., Xu, W.W., Gerhardt, D.J., Albert, T.J., Richmond, T., Jeddloh,
431 J.A., Jia, G., Springer, N.M., Vance, C.P., Stupar, R.M., 2011. The composition and
432 origins of genomic variation among individuals of the soybean reference cultivar
433 Williams 82. Plant Physiol. 155, 645–55.
- 434 Hollox, E.J., Huffmeier, U., Zeeuwen, P.L., Palla, R., Lascorz, J., Rodijk-Olthuis, D., van de
435 Kerkhof, P.C., Traupe, H., de Jongh, G., den Heijer, M., Reis, A., Armour, J.A.,
436 Schalkwijk, J., 2008. Psoriasis is associated with increased beta-defensin genomic copy
437 number. Nat Genet. 40, 23–5.
- 438 Jang, J.S., Simon, V.A., Feddersen, R.M., Rakhshan, F., Schultz, D.A., Zschunke, M.A.,
439 Lingle, W.L., Kolbert, C.P., Jen, J., 2011. Quantitative miRNA expression analysis using
440 fluidigm microfluidics dynamic arrays. BMC Genomics 12, 144.

- 441 Kim, W.-J., Jung, H., Gaffney, P.M., 2011. Development of type I genetic markers from
442 expressed sequence tags in highly polymorphic species. *Mar. Biotechnol. (NY)*. 13, 127–
443 32.
- 444 Linzmeier, R.M., Ganz, T., 2005. Human defensin gene copy number polymorphisms:
445 comprehensive analysis of independent variation in alpha- and beta-defensin regions at
446 8p22-p23. *Genomics* 86, 423–30.
- 447 Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time
448 quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* 25, 402–8.
- 449 Marroni, F., Pinosio, S., Morgante, M., 2014. Structural variation and genome complexity: is
450 dispensable really dispensable? *Curr. Opin. Plant Biol.* 18, 31–6.
- 451 O’Leary, N.A., Gross, P.S., 2006. Genomic structure and transcriptional regulation of the
452 penaeidin gene family from *Litopenaeus vannamei*. *Gene* 371, 75–83.
- 453 Reece, K.S., Ribeiro, W.L., Gaffney, P.M., Carnegie, R.B., Allen, S.K., 2004. Microsatellite
454 marker development and analysis in the eastern oyster (*Crassostrea virginica*):
455 confirmation of null alleles and non-Mendelian segregation ratios. *J. Hered.* 95, 346–52.
- 456 Rosa, R.D., de Lorgeril, J., Tailliez, P., Bruno, R., Piquemal, D., Bachère, E., 2012. A
457 hemocyte gene expression signature correlated with predictive capacity of oysters to
458 survive *Vibrio* infections. *BMC Genomics* 13, 252.
- 459 Rosa, R.D., Santini, A., Fievet, J., Bulet, P., Destoumieux-Garzón, D., Bachère, E., 2011. Big
460 defensins, a diverse family of antimicrobial peptides that follows different patterns of
461 expression in hemocytes of the oyster *Crassostrea gigas*. *PLoS One* 6, e25594.

- 462 Rosa, R.D., Vergnes, A., de Lorgeril, J., Goncalves, P., Perazzolo, L.M., Sauné, L.,
463 Romestand, B., Fievet, J., Gueguen, Y., Bachère, E., Destoumieux-Garzón, D., 2013.
464 Functional divergence in shrimp anti-lipopolysaccharide factors (ALFs): from
465 recognition of cell wall components to antimicrobial activity. PLoS One 8, e67937.
- 466 Saavedra, C., Bachère, E., 2006. Bivalve genomics. Aquaculture 256, 1–14.
- 467 Sauvage, C., Bierne, N., Lapègue, S., Boudry, P., 2007. Single Nucleotide polymorphisms
468 and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. Gene
469 406, 13–22.
- 470 Schmitt, P., Gueguen, Y., Desmarais, E., Bachère, E., de Lorgeril, J., 2010. Molecular
471 diversity of antimicrobial effectors in the oyster *Crassostrea gigas*. BMC Evol. Biol. 10,
472 23.
- 473 Schmitt, P., Rosa, R.D., Duperthuy, M., de Lorgeril, J., Bachère, E., Destoumieux-Garzón, D.,
474 2012. The antimicrobial defense of the Pacific oyster, *Crassostrea gigas*. How diversity
475 may compensate for scarcity in the regulation of resident/pathogenic microflora. Front.
476 Microbiol. 3, 160.
- 477 Schmitt, P., Santini, A., Vergnes, A., Degremont, L., de Lorgeril, J., 2013. Sequence
478 polymorphism and expression variability of *Crassostrea gigas* immune related genes
479 discriminate two oyster lines contrasted in term of resistance to summer mortalities.
480 PLoS One 8, e75900.
- 481 Springer, N.M., Ying, K., Fu, Y., Ji, T., Yeh, C.-T., Jia, Y., Wu, W., Richmond, T., Kitzman,
482 J., Rosenbaum, H., Iniguez, A.L., Barbazuk, W.B., Jeddelloh, J.A., Nettleton, D.,

- 483 Schnable, P.S., 2009. Maize inbreds exhibit high levels of copy number variation (CNV)
484 and presence/absence variation (PAV) in genome content. *PLoS Genet.* 5, e1000734.
- 485 Swanson-Wagner, R.A., Eichten, S.R., Kumari, S., Tiffin, P., Stein, J.C., Ware, D., Springer,
486 N.M., 2010. Pervasive gene content variation and copy number variation in maize and its
487 undomesticated progenitor. *Genome Res.* 20, 1689–99.
- 488 Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics
489 Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596–9.
- 490 Taris, N., Lang, R.P., Camara, M.D., 2008. Sequence polymorphism can produce serious
491 artefacts in real-time PCR assays: hard lessons from Pacific oysters. *BMC Genomics* 9,
492 234.
- 493 Taudien, S., Galgoczy, P., Huse, K., Reichwald, K., Schilhabel, M., Szafranski, K., Shimizu,
494 A., Asakawa, S., Frankish, A., Loncarevic, I.F., Shimizu, N., Siddiqui, R., Platzer, M.,
495 2004. Polymorphic segmental duplications at 8p23.1 challenge the determination of
496 individual defensin gene repertoires and the assembly of a contiguous human reference
497 sequence. *BMC Genomics* 5, 92.
- 498 Wang, Y., Lu, J., Chen, S., Shu, L., Palmer, R.G., Xing, G., Li, Y., Yang, S., Yu, D., Zhao,
499 T., Gai, J., 2014. Exploration of presence/absence variation and corresponding
500 polymorphic markers in soybean genome. *J. Integr. Plant Biol.* 56, 1009–19.
- 501 Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H.,
502 Xiong, Z., Que, H., Xie, Y., Holland, P.W.H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang,
503 J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng,
504 Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Lošo, T., Du, Y., Sun, X.,

505 Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T.,
506 Wang, B., Zhang, J., Peng, Z., Li, Y., Li, N., Wang, J., Chen, M., He, Y., Tan, F., Song,
507 X., Zheng, Q., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang,
508 S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni,
509 P., Miao, G., Wang, J., Wang, Q., Steinberg, C.E.W., Wang, H., Li, N., Qian, L., Zhang,
510 G., Li, Y., Yang, H., Liu, X., Wang, J., Yin, Y., Wang, J., 2012. The oyster genome
511 reveals stress adaptation and complexity of shell formation. *Nature* 490, 49–54.

512 Zhang, L.-M., Luo, H., Liu, Z.-Q., Zhao, Y., Luo, J.-C., Hao, D.-Y., Jing, H.-C., 2014.
513 Genome-wide patterns of large-size presence/absence variants in sorghum. *J. Integr.*
514 *Plant Biol.* 56, 24–37.

515

516

517 **Figure captions**

518 **Figure 1.** Polymorphism in *Cg-BigDef* basal mRNA expression in 163 individual oysters.
519 Relative mRNA levels were analyzed in oysters expressing *Cg-BigDef1* (81/163), *Cg-BigDef2*
520 (47/163) and *Cg-BigDef3* (21/163). Relative expression data were normalized with the Cq
521 geometric mean value of three reference genes (*Cg-ef1a*, *Cg-rpl40* and *Cg-rps6*) according to
522 the $2^{-\Delta\Delta Cq}$ method.

523

524 **Figure 2.** The lack of detection of *Cg-BigDef* expression is likely due to the lack of the
525 corresponding gene in the oyster genome. (A) Relative mRNA abundance of the three *Cg-*
526 *BigDefs* in 10 individual oysters by RT-qPCR. Relative mRNA expressions were normalized
527 with the Cq geometric mean value of three reference genes (*Cg-ef1a*, *Cg-rpl40* and *Cg-rps6*)
528 according to the $2^{-\Delta\Delta Cq}$ method. (B) Detection of *Cg-bigdef* genes by conventional PCR using

529 specific (*Cg-bigdef1*, *Cg-bigdef2*, *Cg-bigdef3*) and universal (*Cg-bigdefs*) primer pairs. The
530 quality and integrity of the gDNA samples were assessed by the amplification of the *Cg-ef1a*
531 gene.

532

533 **Figure 3.** Each *Cg-BigDef* form is encoded by a separated gene. (A) A not-to-scale
534 representation of *Cg-bigdef* genes. Black lines and boxes represent introns and exons,
535 respectively. E=exon. (B) Multiple nucleotide alignment of the three *Cg-bigdef* genes. Exon
536 sequences are in bold. The canonical GT/AG splicing recognition sites are shadowed with
537 black backgrounds and the microsatellite sequences (CTAT, CT and CA) are indicated by
538 rectangles.

539

540 **Figure 4.** Nucleotide alignment of the first exon and intron sequences from alleles of *Cg-*
541 *bigdef1* and *Cg-bigdef2* genes from a same individual oyster. Differences (nucleotide
542 substitutions or gaps) between the two *Cg-bigdef1* alleles and between *Cg-bigdef1* and *Cg-*
543 *bigdef2* genes are shadowed with black and grey backgrounds, respectively. Exon sequences
544 are in bold and the canonical GT/AG splice site junctions are indicated by rectangles.

545

546 **Figure 5.** Phylogeny of the oyster big defensin family. Phylogenetic trees were constructed
547 based on (A) genomic and (B) cDNA sequences using the Neighbour-Joining method in
548 MEGA 4. Bootstrap sampling was reiterated 1,000 times. Sequences included in these
549 analyses were the following: (i) *Cg-bigdef* genes (GenBank: JF703155-JF703160, JN251122,
550 JN251125, JN251126, JN251128, JN251131 and JN251132) and *Cg-defm* gene (GenBank:
551 AM050547) as outgroup; (ii) *Cg-BigDef* transcripts (GenBank: JF703137-JF703154,
552 JN251121, JN251123, JN251124, JN251127, JN251129 and JN251130) and *Cg-Defh2*
553 (GenBank: DQ400102) as outgroup.

554

555 **Supplementary data**

556 **Appendix A.** Cq values of three reference genes (*Cg-ef1 α* , *Cg-rpl40* and *Cg-rps6*) and *Cg-*
557 *BigDefs* (*Cg-BigDef1*, *Cg-BigDef2* and *Cg-BigDef3*) analyzed by a high-throughput RT-
558 qPCR device in 163 individual oysters.

Accepted Manuscript

559 **Tables**560 **Table 1.** Nucleotide sequence of primers.

Primer name	Forward primer (5'-3')	Reverse primer (5'-3')
Primers for RT-qPCR and gDNA amplification		
<i>Cg-ef1a</i>	GAGCGTGAACGTGGTATCAC	ACAGCACAGTCAGCCTGTGA
<i>Cg-rps6</i>	CAGAAGTGCCAGCTGACAGTC	AGAAGCAATCTCACACGGAC
<i>Cg-rpl40</i>	AATCTTGCACCGTCATGCAG	AATCAATCTCTGCTGATCTGG
<i>Cg-bigdef1</i>	TTCGCCTGCTTCCATACTGG	GTCATGGTCACTCCTTATTC
	TCTTTGCTGCTTTGGTGACC	GTCATGGTCACTCCTTATTC
<i>Cg-bigdef2</i>	TCTTTGCTGCTTTGGTGACC	CATAGTTTATCCCCTCCGTC
	TTCGCCTGCTTCCATACTGG	AATGACTGTCATGGTCAGAA
<i>Cg-bigdef3</i>	AGAAGAAGGTGAGACGAG	TGATCCGCACACACCAAACG
	AGAAGAAGGTGAGACGAG	GGCTGATTAATCCATGCAAG C
<i>Cg-bigdef5</i> (universal)	CAGTCTTTGCTGCTTTGGTG	TTGGCACACGAATGACTGTC
	TTCGCCTGCTTCCATACTGG	TTGGCACACGAATGACTGTC
Primers for molecular cloning and sequencing		
<i>Cg-bigdef1</i>	ACGTATAGGACTATCATGGAG	GTCATGGTCACTCCTTATTC
	TTCCTACCAAAGAATAAGGAG	CGTGGCTCAGTAATCTCTGC
<i>Cg-bigdef2</i>	TGTTAACGTATAGGACTATC	AATGACTGTCATGGTCAGAA
	TCTTTGCTGCTTTGGTGACC	CGTGGCTCAGTAATCTCTGC
<i>Cg-bigdef3</i>	TGTTAACGTATAGGACTATC	CGTGGCTCAGTAATCTCTGC
	GGAGAACTGTTAGTGTGCTG	GGCTGATTAATCCATGCAAG C

561

562

563

564 **Table 2.** Occurrence of basal *Cg-BigDef* expression in the oyster population analyzed.

<i>Cg-BigDef</i>	Occurrence (%)	
<i>Cg-BigDef1</i> alone	60	(36.8%)
<i>Cg-BigDef2</i> alone	31	(19%)
<i>Cg-BigDef3</i> alone	08	(4.9%)
<i>Cg-BigDef1</i> and <i>Cg-BigDef2</i>	12	(7.4%)
<i>Cg-BigDef1</i> and <i>Cg-BigDef3</i>	09	(5.5%)
<i>Cg-BigDef2</i> and <i>Cg-BigDef3</i>	04	(2.5%)
All three <i>Cg-BigDef</i> forms	0	(0%)
Absence of <i>Cg-BigDefs</i>	39	(23.9%)
Total of individuals	163	

565

566

567 **Table 3.** Occurrence of basal gene expression of *Cg-BigDefs* according to oyster survival

568 (surviving *versus* non-surviving) to virulent *Vibrio* infections

<i>Cg-BigDef</i>	Occurrence (%)			
	Surviving		Non-surviving	
<i>Cg-BigDef1</i> alone	25	(55.5%)	25	(55.5%)
<i>Cg-BigDef2</i> alone	10	(22.2%)	12	(26.6%)
<i>Cg-BigDef3</i> alone	04	(8.8%)	02	(4.4%)
<i>Cg-BigDef1</i> and <i>Cg-BigDef2</i>	27	(60.0%)	04	(8.8%)
<i>Cg-BigDef1</i> and <i>Cg-BigDef3</i>	06	(13.3%)	31	(68.8%)
<i>Cg-BigDef2</i> and <i>Cg-BigDef3</i>	0	(0%)	0	(0%)
All three <i>Cg-BigDef</i> forms	0	(0%)	0	(0%)
Absence of <i>Cg-BigDefs</i>	12	(26.6%)	10	(22.2%)
Total of individuals	45		45	

569

570

571

Figure 1

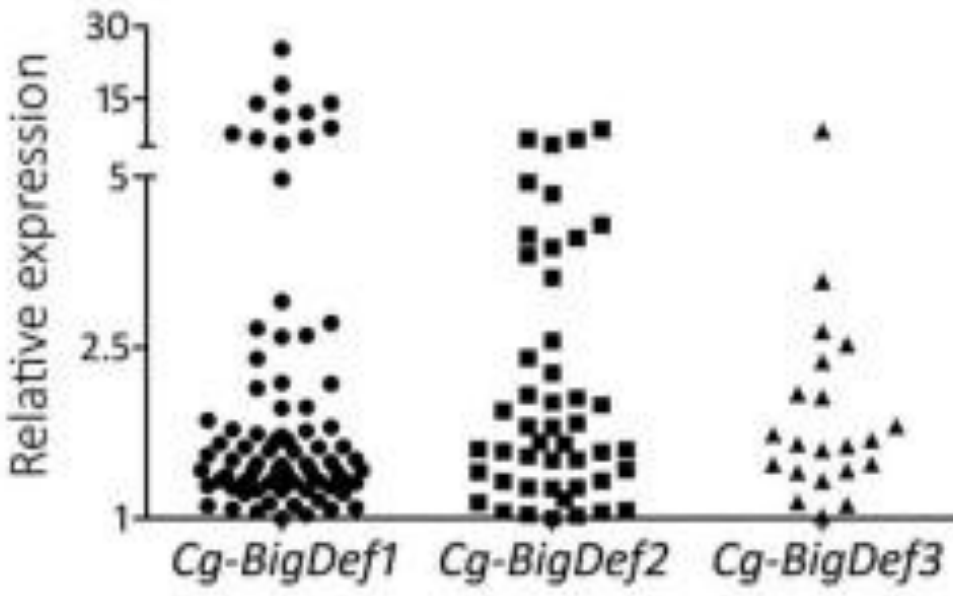


Figure 2

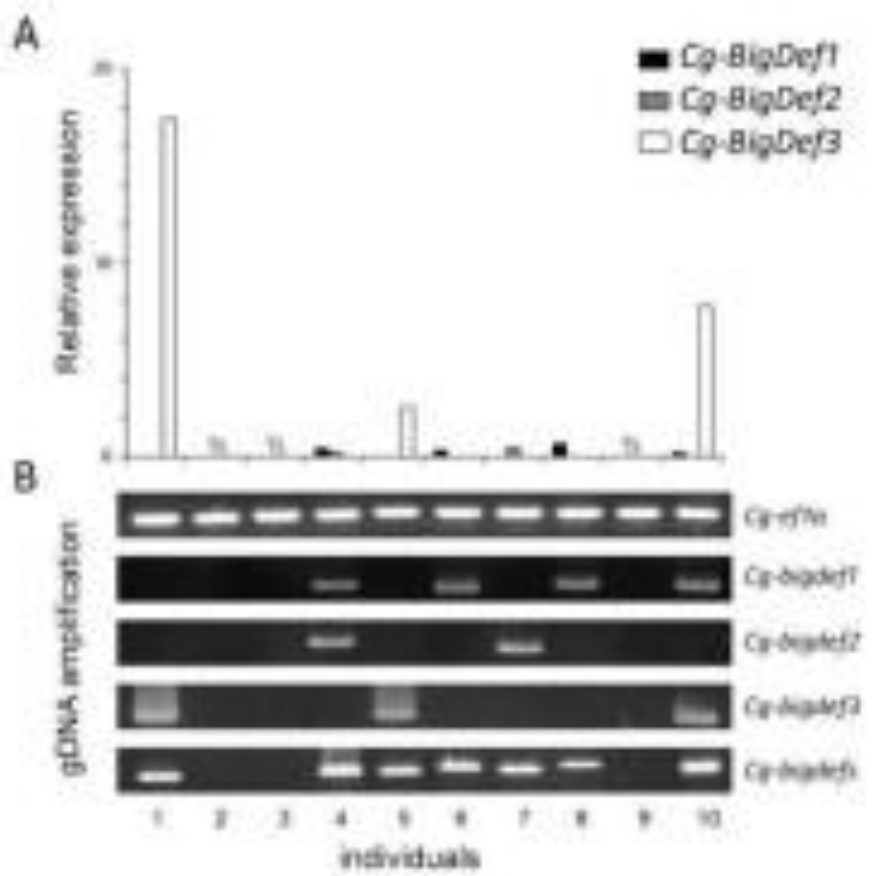


Figure 3

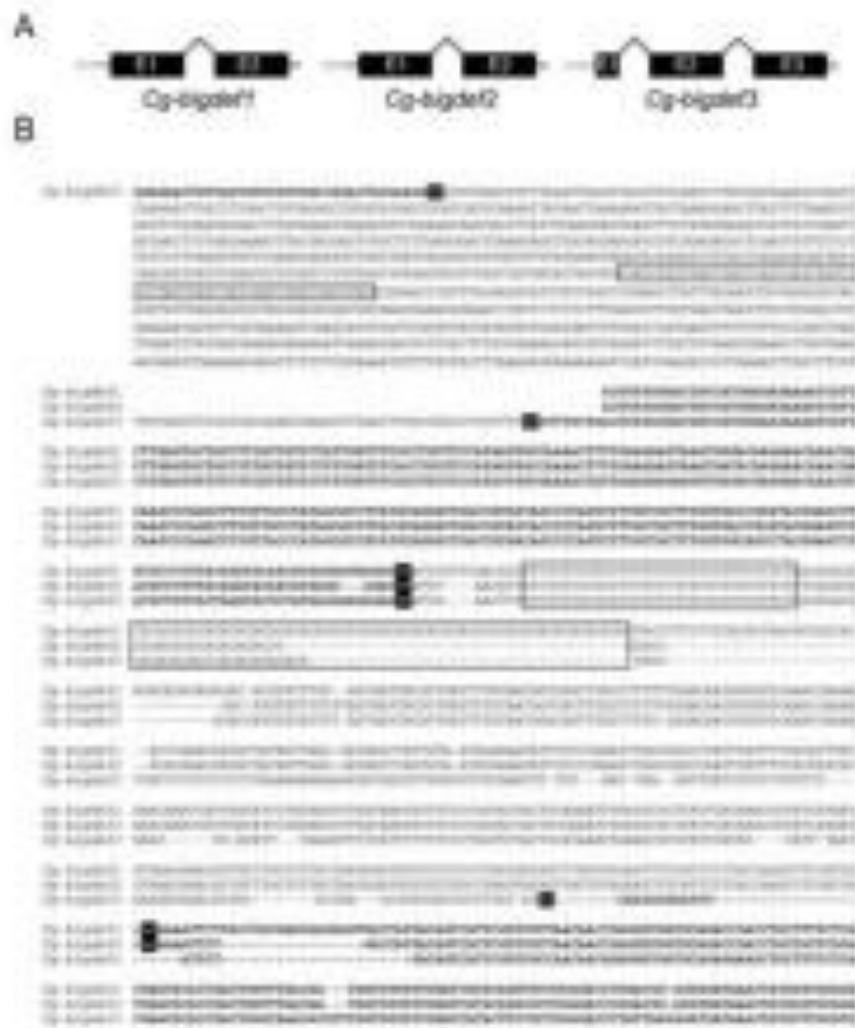


Figure 4

Qy-41y6n F1-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F1-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F1-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F1-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F1-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-1	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT
Qy-41y6n F2-2	AGCTGASGACTCTTCTGGAGAGAGAAAGCTTTTGAGGAGGCTA/CCTTAATGCTTCTTGGCTTCCGGCTTCCGAAAT

Figure 5

