

## Marine Biotechnology

June 2010, Volume 12, Number 3, Pages 326-339

<http://dx.doi.org/10.1007/s10126-009-9227-9>

© 2010 Springer. Part of Springer Science+Business Media

Archimer  
<http://archimer.ifremer.fr>

The original publication is available at <http://www.springerlink.com>

# Microarray-Based Identification of Gonad Transcripts Differentially Expressed Between Lines of Pacific Oyster Selected to Be Resistant or Susceptible to Summer Mortality

Elodie Fleury<sup>1</sup>, Jeanne Moal<sup>1</sup>, Viviane Boulo<sup>2</sup>, Jean-Yves Daniel<sup>1</sup>, David Mazurais<sup>3</sup>, Alain Hénaut<sup>4</sup>, Charlotte Corporeau<sup>1</sup>, Pierre Boudry<sup>1</sup>, Pascal Favrel<sup>1</sup> and Arnaud Huvet<sup>1,\*</sup>

<sup>1</sup> UMR M100 Ifremer- Université de Caen Basse Normandie "Physiologie et Ecophysiologie des Mollusques Marins", Centre de Brest, B.P. 70, 29280 Plouzané/IBFA, IFR ICORE 146, 14032 Caen, France

<sup>2</sup> Ifremer, CNRS Université de Montpellier 2, UMR 5119 ECOLAG CC 80, Place Eugène Bataillon, 34095 Montpellier, France

<sup>3</sup> Ifremer, UMR 1067 Nutrition, aquaculture, et génomique, Centre de Brest, B.P. 70, 29280 Plouzané, France

<sup>4</sup> CNRS Université Pierre et Marie Curie, UMR 7138 Systématique, adaptation, évolution, Case 05, 75252 Paris, France

\*: Corresponding author : A. Huvet, email address : [Arnaud.Huvet@ifremer.fr](mailto:Arnaud.Huvet@ifremer.fr)

## Abstract:

Summer mortality of the Pacific oyster *Crassostrea gigas* is the result of a complex interaction between oysters, their environment, and pathogens. Heredity appears to be a major factor determining the sensitivity of oysters to summer mortality, allowing resistant (R) and susceptible (S) lines to be produced. We conducted genome-wide expression profiling of R and S gonads during the 3-month period preceding a summer mortality event, using a cDNA microarray that we designed. ANOVA analysis revealed that 34 genes were differentially expressed between R and S lines on four dates preceding the mortality event. Annotation of some of these genes highlights reproduction and its allocation and antioxidant defenses as the main pathways that operate differentially between R and S lines. This transcriptional analysis provides new indications to define markers for quantitative trait loci searches and functional studies and evaluate the potential role of each gene in the resistance to summer mortality.

**Keywords:** Mollusca - Bivalvia - Gene expression - cDNA microarrays - Reproduction - Oxidative stress - Energy metabolism

## 1. Introduction

---

Significant mortality of the Pacific oyster has been reported during the summer months in several countries and is of major concern to oyster farmers (Samain and McCombie, 2008). Studies performed within the framework of the multidisciplinary project “Morest” suggested that summer mortality in *Crassostrea gigas* is caused by a complex interaction between oysters, their environment and pathogens (Samain et al., 2007). A large genetic basis was demonstrated for the observed variation in resistance to summer mortality, as 45% of the variance in mortality recorded in the field was observed between families (Degrémont et al., 2005). A high heritability was also estimated for resistance to summer mortality (Degrémont et al., 2007), which offered the possibility to develop lines of oysters that were ‘Resistant’ (R) or ‘Susceptible’ (S) to summer mortality (Samain et al., 2007).

The physiological status of an oyster is suspected to play a role in its resistance to summer mortality. Reproduction, energy metabolism and stress defenses are likely to be the main components of this status (Samain et al., 2007). Due to the high reproductive allocation of oysters, gametogenesis, which occurs in summer, is a period of negative energy budget when most energy is used to produce gametes (Berthelin et al., 2000). Interestingly, R and S oysters made a different investment in reproduction, with R families displaying lower reproductive effort than S families (Samain et al., 2007). This pattern is in agreement with the trade-offs previously suggested between survival and reproduction, meaning that reproductive effort improves at the expense of survival (Ernande et al., 2004). These results suggest that R lines may survive better because they are less reproductively active than S lines (Samain and McCombie, 2008).

Moreover, R and S lines behave differently during the critical period at the start of summer with respect to carbohydrate energy fluxes and their management (Samain et al., 2007), suggesting that R oysters have an increased potential for energy production compared with S oysters. While more energy is produced by R oysters, more of the available energy in the S oysters is directed to reproduction (Samain et al., 2007). This means that more energy is available for the maintenance compartment in R oysters than in S oysters during the critical summer period. This difference has repercussions in terms of capacity for defense against pathogenic agents (e.g. *Vibrio sp* or ostreid herpes virus type 1 OsHV-1) present in the environment and ability to withstand stressful environmental modifications (hypoxia, pollutants, etc.), which are reflected in the differential survival observed between R and S.

To further investigate the nature of the difference between the R and S lines and obtain molecular markers of the resistance to summer mortality, the present study aimed to identify genes differentially expressed between the two lines by genome-wide expression profiling with a cDNA microarray. To date, only one other cDNA microarray has been produced for *Crassostrea*; this contained 4460 sequences derived from *C. virginica* and 2320 from *C. gigas* (Jenny et al., 2007) and was used to study the response of selected families of *C. gigas* to heat stress challenge (Lang et al., 2009). In the present study, within the Network of Excellence ‘Marine Genomics Europe’ (<http://www.marine-genomics-europe.org/>) and the EU funded project ‘Aquafirst’ (<http://www.sigenae.org/aquafirst/>), a much larger cDNA microarray containing 9058 unigenes of *C. gigas* was produced and used to conduct the genome-wide expression profile of gonads sampled from R and S lines on four dates preceding a summer mortality event. Statistical analyses allowed us to pinpoint 34 genes differentially expressed between R and S lines. Candidate genes were prioritized from the literature, based on known biological processes previously suspected to be involved in the mortality phenomenon and for their consistent differences between R and S lines. These candidates were studied by real-time PCR on gonads of R and S oysters of different age, generation and growing environment, sampled just before a mortality event. We discuss the value of these candidate genes for improving knowledge of the genetic factors underlying the susceptibility of the S oysters to summer mortality.

## 2. Materials and methods

---

### 1. Animals sampling

Following a divergent selection experiment on spat survival, a fourth generation of *Crassostrea gigas* lines Resistant (R) and Susceptible (S) to summer mortality was produced in March 2004 at the Ifremer hatchery in La Tremblade (France). Intra-line crosses of “G3c2” generation gave the “G4R” and “G4S” batches according to the crossing designs detailed in Table 1. Briefly, an average of 25 females and 25 males were used for each parental group and crosses were made in both possible directions (male x females and females x males). Gametes from each cross were fecundated separately (12 biparental families) and pools R and S were realized 1 hour post fecundation by mixing in equal proportion the different families. At the end of the larval rearing period, each family was settled on cultch (ground oyster shell). The spat were transferred to the Bouin nursery (France) for on-growing nursery. 500 oysters per G4R and G4S batch were then kept at the Bouin nursery away from mortality risks. At 12 months (March 2005), they were deployed in the field at Fort Espagnol (South Brittany, France). The oysters were cultured in triplicate bags fixed on racks, following local practices. To monitor mortality, all oysters were checked and counted on April 11; May 9, 16, 25 and 30; June 6 and 20; July 7 and 20; and finally August 22. Dead oysters were removed from the bags, and surviving ones returned to the same rearing conditions.

For microarray analysis, R and S oysters were sampled five times (dates 1 to 5: May 9, May 25, June 6, June 20 and July 20, respectively; Figure 1). On each date, 24 oysters per line were randomly collected (8 per bag) and their gonads were immediately dissected. To ensure that each individual will contribute equally to the pool, a piece of 30 mg of gonad tissue was sampled per oyster, pooled (3 pools of 8 oysters), homogenized in Extract-all (Eurobio) at concentration of 1ml/50mg tissue and stored at  $-80^{\circ}\text{C}$  for total RNA extraction. Furthermore, the entire tissues of 10 wild oysters were collected, pooled and homogenized in Extract-all (Eurobio) to constitute a single total RNA sample for use as a reference in all slide hybridizations and RT-PCR analysis.

Four other experiments already done on R and S oyster lines provided RNA for RT-PCR (Fleury et al., 2008) : (1) spat oysters, 4 months old, generation 2, nursery (G2R and G2S); (2) spat oysters, 4 months old, generation 3, nursery (G3R and G3S); (3) adult oysters, 16 months old, generation 2, Rivièrè d’Auray field site, South Brittany (G2 RA R and G2 RA S); (4) adult oysters, 16 months old, generation 2, Baie de Veys field site, Normandy (G2 BDV R and G2 BDV S). In each experiment, the mortality peak was observed one or two weeks after sampling. Following the same procedure used for microarray analysis, 3 pools of 4 gonads were made per line (R and S) in each experiment and prepared for total RNA extraction for RT-PCR analysis.

### 2. RNA preparation

Samples of gonad tissue in Extract-all (Eurobio) were taken frozen and solubilized using a Brinkman Polytron tissue disruptor. Total RNA was then isolated using Extract-all (Eurobio) procedure; the procedure was always performed by the same person. RNA quality was assessed using RNA nano chips and Agilent RNA 6000 nano reagents (Agilent Technologies, Waldbronn, Germany) according to manufacturer’s instructions. RNA concentrations were measured using an ND-1000 spectrophotometer (Nanodrop Technologies) and the samples were stored at  $-80^{\circ}\text{C}$  until use.

For microarray hybridizations, 5 $\mu\text{g}$  of total RNA were directly labeled by reverse transcription and then purified using the Direct ShipShot Labeling kit (Promega), according to manufacturer’s recommendations. This reaction was performed for each of the 24 gonad

samples, with Cy5 (red) incorporation. The reference sample was Cy3-labeled (green) in 24 separate tubes following the same protocol. The 24 Cy3-labeled cDNAs were next pooled, and then divided once more into 24 samples to obtain a homogeneous reference. All dye incorporation rates were verified by ND-1000 spectrophotometer (Nanodrop Technologies). For real-time PCR analysis, RNA samples were first treated with DNase I (1 U/μg total RNA, Sigma) according to the manufacturer's instructions and finally washed twice in 500 μl of Ethanol 95% prior to dilution in 10 μl of RNase/DNase free water. Reverse transcription was then carried out as described by (Huvet et al., 2003) using 4μg of total RNA from each sample.

### 3. cDNA microarray construction

#### cDNA clones and preparation of microarray probes

Sequences from different projects (Marine Genomic Network, (Tanguy et al., 2008) and Aquafirst, Fleury et al., manuscript submitted) together with sequences previously available in Genbank, were assembled in a single database to form 9276 putative transcripts containing 1758 contigs and 7518 singletons (Fleury et al., manuscript submitted, <http://www.sigene.org>). The first step built clusters of sequences sharing at least 75 bp with an identity rate of 96% using MegaBlast (Zhang et al., 2000). The second step constructed coherent contigs from the previous clusters using CAP3 (Huang and Madan, 1999) at the recommended stringency of 40 bp overlap with 90% sequence identity. Considering all cDNA clones, and controls (including water, spotting buffer, and exciting fluorescent targets), a total of 11 088 probes, representing 9058 *Crassostrea gigas* ESTs, were prepared. Inserts were amplified from bacterial templates by PCR (5 min at 95°C; 32 cycles of 0.5 min at 95°C, 0.5 min at 68°C, and 1.5 min at 72°C; a final extension of 4 min at 72°C; PCR system 9600, Applied Biosystems) using universal primers (M13 forward and reverse) located on the cloning vector. Concentrations of the reaction components were: primers at 0.2μM, dNTPs at 0.2mM, 1X Taq buffer (with MgCl<sub>2</sub>), and 0.5 units of Taq DNA polymerase (Qbiogène, Evry, France) in a total volume of 100 μl. PCR products were precipitated with a the solution by 1 volume of isopropanol, redissolved in distilled water, and mixed at 1:1 ratio with printing buffer x2 (20x saline sodium citrate [SSC] diluted to 0.6x SSC in dimethyl sulfoxide [DMSO]). The 20x SSC solution is composed of 3 M NaCl and 0.3 M Na citrate (pH 7.4). The PCR products were examined by running the purified products on a 1% agarose gel, concurrent with a DNA mass ladder. Inserts that failed to amplify (~ 2 %, data not shown) were not replaced. All the produced oyster ESTs and the array design have been submitted to Gene Expression Omnibus (GSE16448).

#### Microarray printing

Microarray slides were printed by the Max Planck Institute (Berlin, Germany) with a GMS417 Arrayer (Genetic Microsystems) on poly-L-lysine-coated slides prepared locally according to a standard protocol. Each slide contained a total of 11 088 features spotted in duplicate to reduce positional bias of the fluorescence readout, and divided up into 48 blocks of 22 rows and 21 columns each. Replicate spots were located within the same block (*i.e.* spotted with the same pin) but were never side-by-side (Additional File 2). Printed slides were exposed to UV light for cross-linking and blocked in a succinic anhydride/sodium borate solution. Excess DNA probe material was then washed off in a water bath at 95°C. Blocked slides were dehydrated in a bath of 100% ethanol, centrifuged, and stored in a dark cool dry location until use.

#### *4. Microarray hybridizations and scanning*

Equimolar amounts of cDNA samples and cDNA reference labeled with Cy5 and Cy3, respectively, were SpeedVac evaporated and mixed into a single pool with the hybridization buffer (ChipHybe™ hybridization buffer, Ventana Discovery, Tucson, AZ, USA). They were then co-hybridized on the same microarray slide, in a Ventana hybridization station (Ventana Discovery, Tucson, AZ, USA). Hybridizations were performed at the INRA IFR 140 transcriptomic facility (Rennes, France). Slides were filled with a pre-hybridization buffer (ChipSpread buffer containing 4×SSC and 0.2% SDS) during 1 hour at 42°C in a humidified chamber. Following this pre-hybridization step, hybridization was conducted overnight at 42°C in a humidified chamber, according to manufacturer's instructions. The arrays were washed twice at room temperature with Ribowash solution (0.1 M Tris, 0.05 M EDTA, and 0.4 M NaCl), twice with 0.1×SSC and finally centrifuged (6000 rpm, 15 seconds, room temperature) for drying.

After the hybridization step, microarray slides were scanned using a Scanner Genepix 4000B (Axon Instruments Inc.) with standard dual laser excitation at 532 nm (17 mW) and 635 nm (10 mW) according to the following parameters: Cy 5 Photo Multiplier Tube (PMT): 550 and Cy 3 PMT: 590. This process was repeated for each of the 24 hybridized slides with the high resolution mode (5 microns resolution). The images (16-bits TIF images) were then analyzed with Genepix pro 5.1 software (Axon Instruments Inc.) according to manufacturer's instructions. Spots were filtered for quality according to the following criteria: diameter of the spot  $\geq 100$ ,  $B_{635} + 2SD > 90$  (background signal and standard deviation of background signal in 635 channel),  $B_{532} + 2SD > 90$  (background signal and standard deviation of background signal in 532 channel),  $F_{635\% \text{ sat}} < 5$  (percent saturated pixels in 635 channel),  $F_{532\% \text{ sat}} < 5$  (percent saturated pixels in 532 channel),  $SNR_{635} > 5$  (signal to noise ratio in the 635 channel) and  $SNR_{532} > 5$  (signal to noise ratio in the 532 channel),  $F_{635} - B_{635} > 400$  (mean fluorescence intensity and background signal in 635) and  $F_{532} - B_{532} > 400$  (mean fluorescence intensity and background signal in 532). Spots not fulfilling these criteria were eliminated.

#### *5. Microarray Data analysis*

##### *5.1 Correction and normalization*

Transformation and normalization of hybridization data were performed to minimize variation arising from technical differences in RNA quality, probe labeling, and hybridization conditions between experiments. As normalization without background subtraction resulted in the lowest variability, microarray data background has not been corrected prior data correction and normalization, as already suggested in Zahurak et al. (2007). First, a logarithmic transformation was performed for each foreground signal intensity (giving "log values"). Considering that expression of the majority of the spotted genes did not change with the experimental conditions, the sums and variances of "log values" should be the same across the different samples. For this reason, the "log values" were median centered as described in (Rezen et al., 2008). Correction was next performed for differences in the variance across the range of gene expression levels, as previously described in Darias et al. (2008) using the formula: (corrected Cy5 log value)<sub>i</sub> = (Cy5 log value)<sub>i</sub> - (Cy3 log value)<sub>i</sub> + (mean Cy3 log value), where Cy5 log value represents sample signal intensity, Cy3 log value represents the reference signal intensity, and mean Cy3 log value represents the mean of all green values obtained for the gene "i" across the different conditions. The corrected and normalized data are available in Additional File 3.

##### *5.2 Identification of differentially expressed genes*

Genes that were differentially expressed between R and S lines were determined statistically ( $p$ -value  $< 1 \times 10^{-4}$ ) by variance analysis using GeneANOVA software (Didier et al., 2002) for the dates in the study. The False Discovery Rate (FDR) associated with the selected genes was thus determined by the formula: [Total number of analyzed genes (9058)  $\times$   $p$ -value ( $1 \times 10^{-4}$ ) / Number of differentially expressed genes]  $\times$  100. For this analysis, we constructed a statistical model with 3 factors: "line" (R or S), "date of sampling" (date 1, 2, 3 and 4) and "technical replicate" (spots deposited in duplicate on the slide). These factors represent a set of interacting parameters reflecting differential gene expression across the experimental conditions. All identified genes were kept for hierarchical clustering analysis (realized with TmeV, <http://www.tm4.org/mev.html>) in order to characterize the expression profiles of statistically relevant genes. Hierarchical clustering is based on average linkage to find the pairs of genes that behave most similarly in all samples. Clusters of interest were identified by visual inspection using TmeV.

### 5.3 Gene ontology analyses

Gene ontology analysis was carried out for all genes spotted on the array, using the BLAST2GO program (Conesa et al., 2005). This program is a Java application enabling gene ontology (GO) assignment, based on data mining of sequences for which no GO annotation is currently available. FASTA-formatted sequences representing all the spotted ESTs were uploaded to the program and BLASTX carried out against the Swiss-Prot database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). GO terms associated with the hits ( $e$ -value  $< 1 \times 10^{-5}$ ) were retrieved by the program and queries were annotated based on hit similarity and GO evidence codes.

### 6. Gene expression analysis by real-time PCR of some selected ESTs

The RNA prepared for microarray analysis and RNA previously sampled in the four other experiments (Fleury et al., 2008) were used for gene expression analysis by real-time PCR of 8 microarray-identified genes selected on the basis of their relevant putative functional annotations. One non-annotated EST was also chosen for real-time PCR analysis based on its very high overall expression levels in both R and S samples. PCR efficiency ( $E$ ) was estimated for each primer pair by determining the slopes of standard curves obtained from serial dilution analysis of the reference cDNA sample (the same than the reference sample Cy3-labeled and used in microarray) to ensure that  $E$  ranged from 98 to 100%. Real-time PCR efficiency ( $E$ ) of one cycle in the exponential phase was calculated according to the equation:  $E = 10^{[-1/\text{slope}]}$  (Pfaffl, 2001). Primer names, accession numbers, PCR efficiencies and sequences are listed in Table 2. In the present study, amplification of elongation factor I (ef I) cDNA (primers in Fabioux et al., 2004) was performed to validate the steady-state level of expression of a housekeeping gene and provide an internal control for gene expression. Indeed, we did not observed any significant difference between Ct values for ef I between the R and S samples ( $t$ -test:  $p = 0.69$ ). Its stability was also assessed by a low coefficient of variation overall the sampling ( $CV = 5.1\%$ ). To date, Siah et al (2008) have also shown the stability of ef I in a marine bivalve species. Real-time PCR amplifications were carried out in triplicate with 5  $\mu$ L cDNA (1/10 dilution) in a total volume of 15  $\mu$ L as described in Huvet et al., 2004 with the iQ SYBR Green Supermix (Biorad) using an Icyler (Biorad). Each run included the cDNA reference, negative controls (each total RNA sample with DNase I treatment), and blank controls (water) analyzed for each primer pair. Melting curve analysis was systematically performed for all genes in order to verify the specificity of the PCR product. For a studied gene "i", results were expressed as  $2^{-\Delta\Delta Ct}$  (Pfaffl, 2001) with  $\Delta Ct = Ct(i) - Ct(\text{ef I})$  and  $\Delta\Delta Ct = \Delta Ct \text{ of cDNA sample} - \Delta Ct \text{ of the cDNA reference}$ . Student  $t$ -tests were realized on the  $\Delta\Delta Ct$  values of each gene between oyster lines Resistant and Susceptible to summer mortality, as proposed in Yuan et al., 2006.

### 3. Results

---

#### *Survival of the oyster lines*

From April until the beginning of June, no mortality was detected in either R or S lines. The first mortalities were observed in mid June, with peak mortality on July 7. At the end of the experiment, 56% and 22% cumulative mortality had been observed for S and R lines respectively, representing a significantly higher mortality rate in S oysters, 2.5 times greater than that of R oysters (Figure 1). R and S oysters sampled on dates 1 to 4 were used in the microarray analysis. Oysters sampled on July 20 (date 5) were not considered; as the only oysters available at this time were those that had survived the mortality event and this would have generated a bias in the sampling strategy (surviving S oysters are supposedly not representative of the “susceptible” character).

#### *Global annotation analysis of spotted ESTs*

Transcriptomic analysis was conducted on R and S oyster lines using the cDNA microarray containing 9058 clones from different oyster libraries. BLAST analyses (with E-value threshold  $< 1 \times 10^{-5}$ ) revealed that 28 % of the spotted clones could be given a functional annotation, 24 % showed hits with databank sequences that had no functional annotation, and 48 % had no hits (data not shown). Gene ontology (GO) analyses of the annotated clones showed that they are involved in more than 1000 distinct biological processes. The variety of these processes can be explained by the nature of the array, which contains cDNA from libraries of different tissues sampled under different conditions. As expected, some of these annotations are related to cellular processes, metabolic processes or biological regulation. The diversity of biological processes represented across this slide will be very helpful for studying the complexity of the summer mortality phenomenon.

#### *Quality and reproducibility of data*

After hybridization and scanning steps, we obtained spots of good quality, with high intensity, homogeneous size and shape and very low background. Indeed, filtering steps (described in the Materials and Methods section) removed less than 1% of the spotted clones. Moreover, the mean correlation coefficient for microarray values, obtained by scatter plot, was 0.94 for technical replicates, and 0.90 for biological replicates.

#### *Identification of differentially expressed genes*

ANOVA analysis revealed that 34 and 2114 genes varied significantly according to the factors “line” and “date of sampling”, respectively, with no interaction between these factors. None of the genes were found to be significantly differentially expressed according to “technical replicate”. The level of significance used to select ESTs as differentially expressed was a  $p$ -value  $< 1 \times 10^{-4}$ . False Discovery Rate (FDR) for the 34 selected ESTs differentially expressed between R and S lines was, therefore, 2.6%.

The ANOVA result is graphically represented in Figure 2, showing genes where a large and statistically significant proportion of variation was due to “line”. Among these 34 selected genes, 16 showed apparent homology to genes of known function in the SwissProt database (E-value  $< 1 \times 10^{-5}$ ). For these annotated ESTs, the top hits corresponding to the BlastX with the lowest E-value (Table 3), were carefully studied and associated publications examined to

extract the most relevant information and annotations using the UniProt Knowledgebase website (<http://www.expasy.ch/sprot/>).

A hierarchical clustering by average linkage was performed to determine whether the genes identified by ANOVA as being differentially expressed exhibited any coordinate patterns of expression between the studied samples. By visual inspection, 2 clusters of coordinately regulated genes were identified (Figure 3), each containing 17 ESTs. The combination of the two types of analysis, annotation studies and clustering, allowed us to identify 17 transcripts over-expressed in S relative to R samples. These were the ESTs annotated as succinate dehydrogenase, tetratricopeptide repeat protein 28, heat shock protein 70kDa, ependymin related protein, CD109 antigen precursor and 12 non-annotated ESTs (Cluster I in Table 3). Conversely, the 17 transcripts identified as over-expressed in R relative to S samples comprised ESTs annotated as catalase, acylcarnitine carrier protein, neuropeptide Y receptor (a member of the diverse group of G-protein coupled receptor), scavenger receptor, MAP-kinase-activated protein, oyster-gonadal-TGF $\beta$ -like, WD repeat-containing protein 92, hexokinase, cytochrome c oxidase, superoxide dismutase, dopamine receptor, and 6 non-annotated ESTs (Cluster II in Table 3).

#### *Validation of microarray data by real-time PCR*

We used two ways to examine the qualitative data produced by the microarray. First, we took advantage of the redundancy of some of the genes included on the slides by comparing the expression data obtained for different ESTs encoding the same genes, *i.e.* calmodulin, myosin and ubiquitin. Indeed for each gene, several ESTs displayed the same best BlastX hit and showed a high degree of similarity (> 90%). These different ESTs encoding calmodulin, myosin and ubiquitin showed similar expression patterns, as detected by the microarray hybridization, with correlation coefficients of 0.92, 0.91 and 0.95, respectively. Second, we assessed the degree of correlation for 9 differentially-expressed genes (Table 3) between the expression data obtained by microarray and the expression data obtained by real-time PCR on the same samples. Eight genes were chosen on the basis of their putative annotations concerning biological processes previously suspected to be important in the mortality phenomenon (reproduction, energy metabolism and defenses) and one non-annotated gene (CK172373) was chosen for its high overall expression levels in both R and S samples. A significant correlation was observed ( $R^2 = 0.89$ ,  $p = 0.002$ ) between the log ratio comparing expression in R and S obtained by microarray and real-time PCR.

#### *Study of 9 differentially expressed genes in several R and S experiments*

To further document the genetic component (*i.e.* the “line” effect), we assayed the level of mRNA of these 9 selected genes by real time PCR in R and S samples collected in 4 other experiments representing different generations, ages and environments (sampling described in Fleury et al., 2008). Similar up and down regulation between R and S was observed for the genes tested in all 4 experiments and in comparison with the data obtained with the microarray data, although differences were not always significant (Figure 4). More precisely, results showed that ESTs corresponding to dopamine receptor, neuropeptide Y receptor and “unknown EST” (CK172373) were systematically differentially expressed between R and S for these 4 experiments in addition to the microarray experiment. Superoxide dismutase, CD-109, acylcarnitine carrier protein and oyster-gonadal-TGF $\beta$ -like were significantly differentially expressed in the two experiments analyzed by real time PCR where significant mortality was observed (1 and 2). Indeed, mortality rates for R and S lines were 4% and 74% , respectively, for 4-month-old G2 oysters (experiment 1) and 30% and 80% for 4-month-old G3 oysters (experiment 2) (Fleury et al., 2008). In 16-month-old oysters, mortality appeared



much lower, with 5% and 14% mortality for R and S, respectively, in experiment 3, and 8% and 5% for experiment 4. Finally, ESTs encoding heat shock protein 70 (HSP70) and cytochrome c oxidase (COX) were each only differentially expressed between R and S lines in one out of the 4 studied experiments (experiments 2 and 4 for HSP70 and COX, respectively) in addition to the microarray experiment.

## 4. Discussion

---

To search for molecular markers linked to the resistance of oyster to summer mortality, we performed a transcriptional analysis to compare resistant and susceptible selected lines. This work consisted of a genome-wide expression profiling of R and S gonads during a temporal survey preceding a mortality event in the field. We hypothesized that the differentially expressed genes revealed by this method could trace the effects of the divergent survival-based selection of the lines and thus represent summer mortality resistance markers even if the structure and function of proteins, because of small variations in genetic sequence, can also affect summer resistance. The present rearing method and survey were performed in a similar way to the work done to obtain the lines by divergent selection (Degrémont et al., 2005). We selected gonad tissue for this study because reproduction is supposedly heavily implicated in *C. gigas* summer mortality (Samain et al., 2007).

Thirty-four genes differentially expressed between oyster lines Resistant and Susceptible to summer mortality

A total of 34 differentially expressed ESTs were identified as being only “line” dependent ( $p$ -value  $< 1 \times 10^{-4}$ ). This low number is similar in magnitude to those obtained in several other experiments that have compared genetically resistant or susceptible animals, e.g., in duodenum tissue of sheep genetically resistant and susceptible to gastrointestinal nematodes (Keane et al., 2006) or in spleen tissue of chickens resistant and susceptible to Marek’s disease (Sarson et al., 2008). This suggests that only a few genes contribute to the difference in survival by interacting with identified environmental risk factors. This is in agreement with the high heritability previously determined for mortality resistance (Samain and McCombie, 2008). The roles of these genes as main genes or modifiers still need to be identified. For example in the nematode, a G-protein coupled receptor, identified as a neuropeptide receptor, is responsible for the difference between solitary and social feeding forms and had subsequently affected the expression of more than 60 other genes (Styer et al., 2008).

In the 34 differentially expressed ESTs, the clustering analysis showed that 17 ESTs were over-expressed and 17 ESTs under-expressed in R compared with S, suggesting that the difference between the two lines is not only due to a physiological weakness of S oysters causing generalized down regulation. Based on this result, we propose to make further expression profiling tests on these 34 ESTs as predictive markers of summer mortality across several years and locations.

The ANOVA realized according to 3 factors (“technical replicate”, “date of sampling” and “line”) showed that no ESTs were differentially expressed according to “technical replicate”. A total of 2114 ESTs varied significantly according to the four sampling dates, independently of the R or S line (see additional file). This high number could be explained by the fact that these sampling dates represented more than one and a half months of active gametogenesis (May to the end of June), during which the oysters underwent large changes in physiological processes associated with gonad development, as shown in trout (Elis et al., 2008). The environmental variations observed in the field during this period (e.g. increase of water temperature, Figure 1) could also be responsible for inducing changes in gene expression (Samain and McCombie, 2008).

The expression patterns of nine of the 34 “line”-dependant genes were further confirmed by real-time PCR on the same samples as those used for microarray analysis. Results showed a significant correlation of the gene expression assays between the two methods ( $R^2 = 0.89$ ).

### *Summer mortality resistance as a selected character*

In order to identify and prioritize genes that could constitute relevant summer mortality resistance markers and candidates of interest for further functional and genetics (search of quantitative trait locus) studies, the expression patterns of these nine selected genes were studied between R and S lines sampled in different conditions. We analyzed R and S samples from four other experiments varying for age (4 or 16 months), generation (second or third generation) and the environment (experimental nursery, or South Brittany or Normandy field sites) with the aim of focusing on the genetic component. Three candidates appeared to be systematically differentially expressed between R and S: dopamine receptor, neuropeptide Y receptor and the non-annotated CK172373. Four other candidates (CD109 antigen precursor, oyster-gonadal-TGF $\beta$ -like, acylcarnitine carrier protein and superoxide dismutase) were found to be significantly differentially expressed between R and S in the two experiments where significant mortality occurred, suggesting an interaction. Heat shock protein 70 and cytochrome c oxidase were considered less important as these candidates were only significantly differentially expressed between R and S in one of the four experiments. Together, these seven selected ESTs constitute preferential candidates for quantitative trait locus analysis and functional studies. Quantitative trait locus linkage studies have recently been started to map regions of the genome influencing resistance to summer mortality (Sauvage, 2008). Once markers linked to quantitative trait loci for resistance to summer mortality have been identified, they can be used in marker-assisted selection. The co-localization of genes identified by transcriptome studies and regions of the genome linked to a studied trait (ie quantitative trait loci) offers a high potential for identifying genes that affect this trait (Darvasi et al., 2003).

What principal pathways operate differently between oyster lines Resistant and Susceptible to summer mortality?

The seven candidate ESTs revealed consistent annotations involved in particular processes suspected to be involved in the summer mortality phenomenon (Samin et al., 2007). For example, the EST corresponding to oyster-gonadal-TGF $\beta$ -like (og-TGF $\beta$ -like), belonging to the Transforming Growth Factor  $\beta$  family, supposedly has a potential role in germinal cell proliferation and/or maturation (Fleury et al., 2008). Based on its over-expression in R compared with S and on the heavy gonad formation in S lines (Samain et al., 2007), a negative effect of og-TGF $\beta$ -like in gonad development was suggested. This has already been observed for some other members of the TGF $\beta$  superfamily with, for example, the involvement of TGF $\beta$ 1 in the inhibition of gonadotropin and 17-alpha, 20-beta-dihydroxyprogesterone(DHP)-induced maturation of zebrafish oocytes (Kohli et al., 2005). Considering that the regulation of reproduction plays a major part in the interaction leading to summer mortality, the identification of a novel G-protein coupled receptor annotated as neuropeptide Y receptor, statistically over expressed in R compared with S, also appears relevant. Indeed, members of this family, through the ligand-receptor interaction (Lecklin et al., 2002), have a strong regulatory influence on reproduction (Magni et al., 2000; Schneider et al., 2004) by down-regulating reproductive investment when energy storage is deficient (Bojkowska et al., 2008). The higher expression of a candidate gene annotated as a G-protein receptor, putatively a neuropeptide Y receptor, in R lines might therefore regulate

energy flow to reduce reproductive investment, e.g. by decreasing egg mass titer as observed in *Lymnea stagnalis* (de Jong-Brink et al., 2001).

The dopamine ligand-receptor system plays a part in the regulation of oxygen assimilation, allocation of energy, growth, reproduction and immune defense (Lacoste et al., 2001a; Lacoste et al., 2001b), through the stimulation of cyclic AMP during stress response (Gotzes et al., 1994). Moreover, its effect on gamete development has been demonstrated in freshwater mussels (Gagne et al., 2003), where dopamine reduces growth of the gonad (Dufour et al., 2005). The higher expression of dopamine receptor observed in R oysters might also reflect differences in reproductive investment observed between R and S lines. The identification of such candidates, involved in reproduction and differentially expressed according to the genetic lines, provides further evidence for the involvement of reproductive investment as one of the factors linked to survival, as has already been suggested in several other invertebrates (Myrand et al., 2000; Partridge et al., 1997). Indeed, it is generally accepted that heavy reproductive investment of animals has a negative effect on survival, through the “reproductive cost” and its repercussions on the energy budget (Browne et al., 1982). For the oyster, trade-offs between survival and reproduction have already been suggested, meaning that reproductive effort improves at the expense of survival (Ernande et al., 2004). ESTs encoding og-TGF $\beta$ -like, NPY receptor and dopamine receptor are therefore starting points for further research on the implication of reproduction in *C. gigas* summer mortality.

In addition to the consequences that reproductive investment has in terms of energy consumption, it is now established that reproduction decreases antioxidant defenses illustrating that reproduction creates an oxidative stress cost (Alonso-Alvarez et al., 2004). Reproductive strategies should influence antioxidant defenses, as already observed in insects (Alonso-Alvarez et al., 2004; Wang et al., 2001). This hypothesis makes the identification of the differentially expressed ESTs annotated as catalase and superoxide dismutase particularly relevant. These are the principal antioxidant enzymes involved in maintaining cellular homeostasis and antioxidant defense by removing Reactive Oxygen Species (ROS), produced under stress (Jo et al., 2008). Moreover, it has been recently shown that superoxide dismutase may participate in the regulation of the sepsis-induced inflammatory response by detoxifying internal oyster tissues or improving the ability of hemocytes to resist oxidant injury (Gonzalez et al., 2005). Indeed, activation and/or deterioration of the electron transport chain in mitochondria is the principal cause of oxidative stress increase, due to excessive production of toxic oxygen radicals. Such effects could be related to chronic stress from chemical pollutants or environmental stressors (Samain and McCombie, 2008), as revealed by the differential expression over the four studied dates from April to June. It can also be linked to reproduction because mitochondria emerged as a highly solicited subcellular compartment during reproduction in bivalves (Kraffe et al., 2008). In these conditions, cells use many antioxidant defense strategies and consume much more energy to control their ROS level. When production of free radicals is excessive, some DNA damage, carbohydrate oxidation (Manduzio et al., 2005) and disturbance of the immune defense system were observed (Gagnaire et al., 2006). Then, a greater capacity of R oysters to detoxify the cell through antioxidant enzymes should play a role in their mortality resistance, as previously demonstrated by (Lambert et al., 2007); this effect would be strengthened here by the over-expression of these antioxidant enzymes in R lines. In vertebrates, stress response stimulates metabolic pathways such as the AMP-dependent protein kinase (AMPK) pathway that turns off energy-consuming processes and turns on those that generate ATP, such as lipid oxidation (Hardie et al., 2003). Indeed, AMPK function is regulated at the post-translational level (Krämer et al., 2007) and AMPK activity stimulates mitochondrial fatty acid oxidation to produce energy (Hardie et al., 1992). In our study, the higher expression of the EST annotated as acylcarnitine carrier protein (ACP) supports the hypothesis of an increased potential of R oysters for energy production through lipid oxidation. Indeed, ACP has the important function of fostering oxidation of long-chain fatty

acids by the mitochondria and stimulating protein-sparing action by increasing the energy derived from lipids (Zhang et al., 2005). Its higher expression level detected in R compared to S oysters should therefore suggest a better optimization of energy production through lipid metabolism.

## **Conclusion**

---

The present study has provided an extensive view of the genetic basis underlying the differences between oyster lines selected for resistance or susceptibility to summer mortality, leading to an emphasis on differences in reproductive activity and allocation, and antioxidant defenses, as the main pathways operating differentially between R and S lines. However, this microarray analysis should be considered as an initial screening for genes of interest that is to be followed by more specific and precise individual investigation of these genes. To make further evaluation of the potential role of each gene in the resistance to summer mortality, this work emphasizes the need for functional studies to unravel the purpose of these genes. The 18 ESTs with no annotations but which exhibited patterns of expression correlated with known genes showing differential expression between R and S also require further consideration. Knowledge of spatio-temporal or stimulus-dependent expression patterns of a given gene is valuable information for assigning a putative function. However, gene function is primarily assessed on the basis of altered phenotypes associated with gene disruption. The development and use of functional analysis of all selected genes, in association with the mapping of these genes, will be one of the starting points for further research to elucidate the basis of summer mortality. We clearly expect that genetic mapping of the identified genes of which the expression can explain summer survival will shed light on the architecture of this trait (see Rockman and Kruglyak, 2006). Finally, the 34 significant candidates should be further profiled to predict the susceptibility of oysters to summer mortality and the occurrence of summer mortality events.

## **Acknowledgements**

---

This study was conducted as part of the European program “Aquafirst” and the network of excellence “Marine Genomic Europe”. The authors are grateful to J. Lafore, V. Quillien, G. Davey and P. Lindeque for their support during the course of this work. We are indebted to Helen McCombie for her help in editing the English language. We also thank P. Prunet, A. Le Cam, S. Mativet, G. Leborgne from INRA-SCRIBE for excellent technical assistance, the SIGENAE team (INRA Toulouse, France) for bioinformatic tool development (<http://www.sigенаe.org>). We acknowledge the Max Planck Institute of Berlin (Germany) for the preparation of the slides. E. Fleury was funded by Ifremer and a Région Basse-Normandie doctoral grant.

## References

---

- Alonso-Alvarez, C., Bertrand, S., Devevey, G., Prost, J., Faivre, B., Sorci, G. (2004) Increased susceptibility to oxidative stress as a proximate cost of reproduction. *Ecology Letters* 7: 363-368.
- Berthelin, C., Kellner, K., Mathieu, M. (2000) Storage metabolism in the Pacific oyster (*Crassostrea gigas*) in relation to summer mortalities and reproductive cycle (west coast of France). *Comp Biochem Physiol B Biochem Mol Biol* 125: 359-369.
- Bojkowska, K., Hamczyk, M.M., Tsai, H.W., Riggan, A., Rissman, E.F. (2008) Neuropeptide Y influences acute food intake and energy status affects NPY immunoreactivity in the female musk shrew (*Suncus murinus*). *Hormones and behavior* 53: 342-350.
- Browne, R.A. (1982) The costs of reproduction in brine shrimp. *Ecology* 63: 43-47.
- Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674-3676.
- Darias, M.J., Zambonino-Infante, J.L., Hugot, K., Cahu, C.L., Mazurais, D. (2008) Gene expression patterns during the larval development of European sea bass (*Dicentrarchus labrax*) by microarray analysis. *Marine Biotechnology* 10: 416-428.
- Darvasi, A. (2003) Genomics: Gene expression meets genetics. *Nature* 422: 269-270.
- De Jong-Brink, M., ter Maat, A., Tensen, C.P. (2001) NPY in invertebrates: molecular answers to altered functions during evolution. *Peptides* 22: 309-315.
- Degrémont, L., Bedier, E., Soletchnik, P., Ropert, M., Huvet, A., Moal, J., Samain, J., Boudry, P. (2005) Relative importance of family, site, and field placement timing on survival, growth, and yield of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). *Aquaculture* 249: 213-229.
- Degrémont, L., Ernande, B., Bédier, E., Boudry, P. (2007) Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for survival and growth. *Aquaculture* 262: 41-53.
- Didier, G., Brezellec, P., Remy, E., Henaut, A. (2002) GeneANOVA--gene expression analysis of variance. *Bioinformatics* 18: 490-491.
- Dufour, S., Weltzien, F.A., Sebert, M.E., Le Belle, N., Vidal, B., Vernier, P., Pasqualini, C. (2005) Dopaminergic inhibition of reproduction in teleost fishes: ecophysiological and evolutionary implications. *Ann N Y Acad Sci* 1040: 9-21.
- Elis, S., Batellier, F., Couty, I., Balzergue, S., Martin-Magniette, M.L., Monget, P., Blesbois, E., Govoroun, M.S. (2008) Search for the genes involved in oocyte maturation and early embryo development in the hen. *BMC Genomics* 9: 110.
- Ernande, B., Boudry, P., Clobert, J., Haure, J. (2004) Plasticity in resource allocation based life history traits in the Pacific oyster, *Crassostrea gigas*. I. Spatial variation in food abundance. *J Evol Biol* 17: 342-356.
- Fabioux, C., Pouvreau, S., Le Roux, F., Huvet, A. (2004) The oyster vasa-like gene: a specific marker of the germline in *Crassostrea gigas*. *Biochem Biophys Res Commun* 315: 897-904.
- Fleury, E., Fabioux, C., Lelong, C., Favrel, P., Huvet, A. (2008) Characterization of a gonad-specific transforming growth factor-beta superfamily member differentially expressed during the reproductive cycle of the oyster *Crassostrea gigas*. *Gene* 410: 187-196.
- Fleury, E., Huvet, A., Lelong, C., de Lorgeril, J., Boulo, V., Gueguen, Y., Bachère, E., Tanguy, A., Moraga, D., Fabioux, C., Lindeque, P., Shaw, J., Reinhardt, R., Prunet, P., Davey, G., Lapegue, S., Sauvage, C., Corporeau, C., Moal, J., Gavory, F., Wincker, P., Moreews, F., Klopp, C., Mathieu, M., Boudry, P., Favrel, P. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics*, submitted.

Gagnaire, B., Thomas-Guyon, H., Burgeot, T., Renault, T. (2006) Pollutant effects on Pacific oyster, *Crassostrea gigas* (Thunberg), hemocytes: screening of 23 molecules using flow cytometry. *Cell Biol Toxicol* 22: 1-14.

Gagne, F., Blaise, C. (2003) Effects of municipal effluents on serotonin and dopamine levels in the freshwater mussel *Elliptio complanata*. *Comp Biochem Physiol C Toxicol Pharmacol* 136: 117-125.

Gonzalez, M., Romestand, B., Fievet, J., Huvet, A., Lebart, M.C., Gueguen, Y., Bachere, E. (2005) Evidence in oyster of a plasma extracellular superoxide dismutase which binds LPS. *Biochem Biophys Res Commun* 338: 1089-1097.

Gotzes, F., Balfanz, S., Baumann, A. (1994) Primary structure and functional characterization of a *Drosophila* dopamine receptor with high homology to human D1/5 receptors. *Receptors Channels* 2: 131-141.

Hardie, D.G. (1992) Regulation of fatty acid and cholesterol metabolism by the AMP-activated protein kinase. *BBA* 1123: 231-238.

Hardie, D.G. (2003) Minireview: the AMP-activated protein kinase cascade: the key sensor of cellular energy status. *Endocrinology* 144: 5179-5183.

Huang, X., Madan, A. (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9: 868-877.

Huvet, A., Daniel, J.Y., Quéré, C., Dubois, S., Prudence, M., Van Wormhoudt, A., Sellos, D., Samain, J.F., Moal, J. (2003) Tissue expression of two alpha-amylase genes in the Pacific oyster *Crassostrea gigas*. Effect of two different food rations. *Aquaculture*: 321-333.

Huvet, A., Herpin, A., Degremont, L., Labreuche, Y., Samain, J.F., Cunningham, C. (2004) The identification of genes from the oyster *Crassostrea gigas* that are differentially expressed in progeny exhibiting opposed susceptibility to summer mortality. *Gene* 343: 211-220.

Jenny, M.J., Chapman, R.W., Mancina, A., Chen, Y.A., McKillen, D.J., Trent, H., Lang, P., Escoubas, J.M., Bachere, E., Boulo, V., Liu, Z.J., Gross, P.S., Cunningham, C., Cupit, P.M., Tanguy, A., Guo, X., Moraga, D., Boutet, I., Huvet, A., De Guise, S., Almeida, J.S., Warr, G.W. (2007) A cDNA microarray for *Crassostrea virginica* and *C. gigas*. *Mar Biotechnol* (NY) 9: 577-591.

Jo, P.G., Choi, Y.K., Choi, C.Y. (2008) Cloning and mRNA expression of antioxidant enzymes in the Pacific oyster, *Crassostrea gigas* in response to cadmium exposure. *Comp Biochem Physiol C Toxicol Pharmacol* 147: 460-469.

Keane, O.M., Zadissa, A., Wilson, T., Hyndman, D.L., Greer, G.J., Baird, D.B., McCulloch, A.F., Crawford, A.M., McEwan, J.C. (2006) Gene expression profiling of naive sheep genetically resistant and susceptible to gastrointestinal nematodes. *BMC Genomics* 7: 42.

Kohli, G., Clelland, E., Peng, C. (2005) Potential targets of transforming growth factor-beta1 during inhibition of oocyte maturation in zebrafish. *Reprod Biol Endocrinol* 3: 53.

Kraffe, E., Tremblay, R., Belvin, S., Le Coz, J. R., Marty, Y., Guderley, H. (2008) Effect of reproduction on escape responses, metabolic rates and muscle mitochondrial properties in the scallop *Placopecten magellanicus*. *Mar. Biol.* 156: 25-38.

Krämer, D.K., Al-Khalili, L., Guigas, B., Leng, Y., Garcia-Roves, P.M., Krook, A. (2007) Role of AMPK and PPARdelta in the regulation of lipid and glucose metabolism in human skeletal muscle. *J Biol Chem.* 282: 19313-19320.

Lacoste, A., Jalabert, F., Malham, S.K., Cueff, A., Poulet, S.A. (2001b) Stress and stress-induced neuroendocrine changes increase the susceptibility of juvenile oysters (*Crassostrea gigas*) to *Vibrio splendidus*. *Appl Environ Microbiol* 67: 2304-2309.

Lacoste, A., Malham, S.A., Cueff, A., Jalabert, F., Gelebart, F., Poulet, S.A. (2001a) Evidence for a form of adrenergic response to stress in the mollusc *Crassostrea gigas*. *Journal of Experimental Biology* 204: 1247-1255.

Lambert, C., Soudant, P., Delaporte, M., Moal, J., Boudry, P., Jean, F., Huvet, A., Samain, J.-F. (2007) Hemocyte characteristics in families of oysters, *Crassostrea gigas*, selected for differential survival during summer and reared in three sites. *Aquaculture* 270: 276-288.

Lang, R.P., Bayne, C.J., Camara, M.D., Cunningham, C., Jenny, M.J., Langdon, C.J. (2009) Transcriptome Profiling of Selectively Bred Pacific Oyster *Crassostrea gigas* Families that Differ in Tolerance of Heat Shock. *Mar Biotechnol* (NY) in press:

- Lecklin, A., Lundell, I., Paananen, L., Wikberg, J.E., Mannisto, P.T., Larhammar, D. (2002) Receptor subtypes Y1 and Y5 mediate neuropeptide Y induced feeding in the guinea-pig. *Br J Pharmacol* 135: 2029-2037.
- Magni, P., Motta, M., Martini, L. (2000) Leptin: a possible link between food intake, energy expenditure, and reproductive function. *Regul Pept* 92: 51-56.
- Manduzio, H., Cosette, P., Gricourt, L., Jouenne, T., Lenz, C., Andersen, O.K., Leboulenger, F., Rocher, B. (2005) Proteome modifications of blue mussel (*Mytilus edulis* L.) gills as an effect of water pollution. *Proteomics* 5: 4958-4963.
- Myrand, B., Guderley, H., Himmelman, J. (2000) Reproduction and summer mortality of blue mussels (*Mytilus edulis* L.) in the Magdalen Islands, southern Gulf of St. Lawrence. *Mar Ecol Prog Ser* 197: 193-207.
- Partridge, L., Prowse, N. (1997) The effects of reproduction on longevity and fertility in male *Drosophila melanogaster*. *J Insect Physiol* 43: 501-512.
- Pfaffl M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29: 2002-2007.
- Rezen, T., Juvan, P., Fon Tacer, K., Kuzman, D., Roth, A., Pompon, D., Aggerbeck, L.P., Meyer, U.A., Rozman, D. (2008) The Sterolgene v0 cDNA microarray: a systemic approach to studies of cholesterol homeostasis and drug metabolism. *BMC Genomics* 9: 76.
- Rockman, M.V., Kruglyak, L. (2006) Genetics of global gene expression. *Nature Reviews Genetics* 7: 862-872.
- Samain, J.F., Dégremont, L., Soletchnik, P., Haure, J., Bédier, E., Ropert, M., Moal, J., Huvet, A., Bacca, H., Van Wormhoudt, A., Delaporte, M., Costil, K., Pouvreau, S., Lambert, C., Boulo, V., Soudant, P., Nicolas, J.L., Le Roux, F., Renault, T., Gagnaire, B., Géret, F., Boutet, I., Burgeot, T., Boudry, P. (2007) Genetically based resistance to summer mortality in the Pacific oyster (*Crassostrea gigas*) and its relationship with physiological, immunological characteristics and infection process. *Aquaculture* 268: 227-243.
- Samain, J.F., McCombie, H., 2008. Summer Mortality of Pacific Oyster *Crassostrea gigas*. Ed. Ifremer/Quae, Versailles, France. 332pp.
- Sarson, A.J., Parvizi, P., Lepp, D., Quinton, M., Sharif, S. (2008) Transcriptional analysis of host responses to Marek's disease virus infection in genetically resistant and susceptible chickens. *Anim Genet* 39: 232-240.
- Sauvage, C. (2008) Développement de marqueurs moléculaires liés à la résistance à la mortalité estivale chez l'huître creuse *Crassostrea gigas*. Approche QTL. Thesis of the La Rochelle University, France. 270pp.
- Schneider, J.E. (2004) Energy balance and reproduction. *Physiol Behav* 81: 289-317.
- Siah, A., Dohoo, C., McKenna, P., Delaporte, M., Berthe, F.C. (2008) Selecting a set of housekeeping genes for quantitative real-time PCR in normal and tetraploid haemocytes of soft-shell clams, *Mya arenaria*. *Fish Shellfish Immunol.* 25: 202-207.
- Styer, K.L., Singh, V., Macosko, E., Steele, S.E., Bargmann, C.I., Aballay, A. (2008) Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* 322: 460-464.
- Tanguy, A., Bierne, N., Saavedra, C., Pina, B., Bachere, E., Kube, M., Bazin, E., Bonhomme, F., Boudry, P., Boulo, V., Boutet, I., Cancela, L., Dossat, C., Favrel, P., Huvet, A., Jarque, S., Jollivet, D., Klages, S., Lapegue, S., Leite, R., Moal, J., Moraga, D., Reinhardt, R., Samain, J.F., Zouros, E., Canario, A. (2008) Increasing genomic information in bivalves through new EST collections in four species: development of new genetic markers for environmental studies and genome evolution. *Gene* 408: 27-36.
- Wang, Y., Salmon, A.B., Harshman, L.G. (2001) A cost of reproduction: oxidative stress susceptibility is associated with increased egg production in *Drosophila melanogaster*. *Exp Gerontol* 36: 1349-1359.
- Yuan, J.S., Reed, A., Chen, F., Stewart, C.N., Jr. (2006) Statistical analysis of real-time PCR data. *BMC Bioinformatics* 7: 85.
- Zahurak, M., Parmigiani, G., Yu, W., Scharpf, R.B., Berman, D., Schaeffer, E., Shabbeer, S., Cope, L. (2007) Pre-processing Agilent microarray data. *BMC Bioinformatics* 8: 142.

Zhang, Z., Schwartz, S., Wagner, L., Miller, W. (2000) A greedy algorithm for aligning DNA sequences. *J Comput Biol* 7: 203-214.

Zhang, D.M., Yoshimatsu, T., Furuse, M. (2005) Effects of L-carnitine enrichment on the population growth, egg ratio and body size of the marine rotifer, *Brachionus rotundiformis*. *Aquaculture* 248: 51-57.

## Tables

Table 1: Crosses to constitute the pools of R (resistant to summer mortality) and S (susceptible to summer mortality) oyster lines. G0 represents the oysters collected from natural settlement in the Marennes-Oléron bay (France), and then G1, G2 and G3 are following generation of production (Samain and McCombie, 2008). In bold: Each G4 batch produced (R above and S below) and then mixed in equal proportion to obtain the R and the S pools.

Pool R			Sire G0				
			9		15		
Family G1			F9-35	F9-36	F15-57	F15-58	
G2C1			J2	O2	R2	W2	
Sire G0	Family G1	G2C1	G3C1	JJ2	OO2	RR2	WW2
2	F2-5	A2	AA2	<b>AAJJ2</b>	<b>AAOO2</b>	<b>AARR2</b>	<b>AAWW2</b>
	F2-8	G2	GG2	<b>GGJJ2</b>	<b>GGOO2</b>	<b>GGRR2</b>	<b>GGWW2</b>
9	F9-35	J2	JJ2			<b>JJRR2</b>	<b>JJWW2</b>
	F9-36	O2	OO2			<b>OORR2</b>	<b>OOWW2</b>

Pool S			Sire G0				
			7		14		
Family G1			F7-25	F7-26	F14-54	F14-54/F14-55	
G2C1			E2	L2	M2	P2	
Sire G0	Family G1	G2C1	G3C1	EE2	LL2	MM2	PP2
4	F4-16/F4-15	B2	BB2	<b>BBEE2</b>	<b>BBLL2</b>	<b>BBMM2</b>	<b>BBPP2</b>
	F4-15/F4-16	Y2	YY2	<b>YYEE2</b>	<b>YYLL2</b>	<b>YYMM2</b>	<b>YYPP2</b>
7	F7-25	E2	EE2			<b>EEMM2</b>	<b>EEPP2</b>
	F7-26	L2	LL2			<b>LLMM2</b>	<b>LLPP2</b>



Table 2: Sense and anti-sense sequences of primer pairs for real-time PCR analysis, used to amplify cDNA of 9 ESTs differentially expressed between Resistant and Susceptible oysters. PCR efficiencies were determined for each primer pair by determining the slopes of standard curves obtained from serial dilution analysis of the cDNA reference.

GenBank #	EST name	Forward primer	Reverse primer	Efficiency (%)
AM855194	Acylcarnitine carrier protein	5'-GTG GCG TGT GTC TGG TAG TC-3'	5'-TCC TTT ATA CAG CGG CGT CT-3'	99.7
AM856249	Neuropeptide Y receptor	5'-GTG GCT TGT GGG CTT ATT GT-3'	5'-CTG AAA TCC GAA TGG ACG AC-3'	101.9
AM856765	Oyster-Gonadal-TGF $\beta$ -like	5'-TTG GAC ATC AGG GAA ATT CTG-3'	5'-CCA AAC GAA ACG ACA GGA AC-3'	98.7
CX068896	Cytochrome c	5'-GCT TAC GCC GGT CTG AAC T-3'	5'-GCT AAA GGG ATT GGC ACC TC-3'	101.6
CU681620	Superoxide dismutase	5'-AGT CTG GTC GCA CAT TCT TGT-3'	5'-CAT GTG CCA ATC AAG ATC CTC-3'	101.4
FP091069	Dopamine receptor	5'-ATC ACG AGT AAG GCG ACG AG-3'	5'-CGG TGT TTG GTA ATG TGC TG-3'	101.3
AM854673	Heat shock 70 kDa	5'-CCT GCA ATA TGG AGT GAT TCG-3'	5'-CTT CCG GTT CAT AAG CCA TC-3'	100.5
CX069254	CD109 antigen precursor	5'-CTG GTT GGC TGG CTC ATA GT-3'	5'-ATT TCC ACC ACT CCA ACC TG-3'	98.8
CK172373	No annotation	5'-ACA TCA GGT TTA CGG CGT TC-3'	5'-TGC CCA CCA ATA ACA ATG C-3'	98.7

Table 3: List of differentially expressed ESTs between Resistant and Susceptible oyster lines, obtained by microarray analysis. Genes subsequently studied by real-time PCR are shown in bold. GenBank accession numbers are indicated. The Sigenae contig name (<http://www.sigenae.org>) used for Blast comparison against the Swiss-Prot database is given. Resulting best hit (Swiss-Prot hit description) and corresponding score (E-Value) are indicated. NA indicates “No available Annotation” in the Swiss-Prot database. P-value ANOVA represents the p-value obtained by GeneANOVA analysis allowing the selection of differentially expressed genes (p-value < 1x10e-4). Cluster I groups ESTs under-expressed in R compared with S, cluster II groups ESTs over-expressed in R compared with S. p-value was obtained by ANOVA analysis.

Genbank #	Sigenae name	Swissprot hit description	E-value	p-value ANOVA	Cluster
FP089729	oyih10b03k06r1_m13r.1.a.cg.1	Succinate dehydrogenase	2.00E-62	1.00E-05	
AM855200	cdn20p0002o22.f.1.a.cg.1	NA		1.00E-06	
AM855201	cdn20p0002o24.f.1.a.cg.1	NA		1.90E-05	
AM858582	cdn37p0004k03.f.1.a.cg.1	NA		2.00E-05	
AM858905	cdn37p0005i17.f.1.a.cg.1	NA		2.10E-06	
AM854479	cdn19p0004j23.f.1.a.cg.1	NA		2.50E-05	
BQ426558	BQ426558.1.a.cg.1	NA		3.20E-05	
BQ426632	BQ426632.1.a.cg.1	NA		3.30E-05	
AM237704	AM237704.1.a.cg.1	NA		3.60E-05	Cluster I
CU682072	oygd09b08j20r1_m13rev.1.a.cg.1	NA		3.90E-05	
CU683250	oyge09b11e21r1_m13rev.1.a.cg.1	NA		4.00E-05	
CU683279	oyge09b11e22r1_m13rev.1.a.cg.1	Tetratricopeptide repeat protein 28	2.00E-25	4.10E-06	
CU684430	oyge10b14l11r1_m13rev.1.a.cg.1	NA		4.20E-05	
CU685075	oypg10b07l21r1_m13rev.1.a.cg.1	NA		5.20E-09	
<b>AM854673</b>	<b>cdn20p0001d16.f.1.a.cg.1</b>	<b>Heat shock 70 kDa protein</b>	<b>5.00E-24</b>	<b>5.40E-05</b>	
CU682651	oygd10b10m11r1_m13rev.1.a.cg.1	Mammalian ependymin-related protein	2.00E-09	5.80E-07	
<b>CB617381</b>	<b>CB617381.1.a.cg.1</b>	<b>CD109 antigen precursor</b>	<b>7.00E-12</b>	<b>6.00E-06</b>	
AM853618	cdn19p0002c05.f.1.a.cg.1	Catalase	2.00E-83	6.00E-05	
<b>AM855194</b>	<b>cdn20p0002o15.f.1.a.cg.1</b>	<b>acylcarnitine carrier protein</b>	<b>1.00E-35</b>	<b>6.20E-05</b>	
AM855858	cdn20p0005g01.f.1.a.cg.1	NA		6.20E-06	
AM856073	cdn21p0001b19.f.1.a.cg.1	NA		6.30E-05	
<b>AM856249</b>	<b>cdn21p0001j08.f.1.a.cg.1</b>	<b>Neuropeptide Y receptor</b>	<b>5.00E-06</b>	<b>6.50E-05</b>	
AM856616	cdn21p0002j12.f.1.a.cg.1	Scavenger receptor	9.00E-08	7.00E-05	
AM856617	cdn21p0002j13.f.1.a.cg.1	MAP kinase-activated protein	1.00E-77	7.30E-05	
<b>AM856765</b>	<b>cdn21p0003a17.f.1.a.cg.1</b>	<b>Oyster-gonadal-TGFβ-like</b>	<b>0</b>	<b>7.40E-05</b>	
AM856836	cdn21p0003d20.f.1.a.cg.1	WD repeat-containing protein 92	6.00E-117	7.80E-05	Cluster II
CU684816	oypg09b05m11r1_m13rev.1.a.cg.1	Hexokinase type I	7.00E-09	8.20E-05	
CU681474	oygd09b07d18r1_m13rev.1.a.cg.1	<b>Cytochrome c oxidase</b>	<b>2.00E-53</b>	<b>8.40E-05</b>	
<b>CK172373</b>	<b>CK172373.1.a.cg.1</b>	<b>NA</b>		<b>9.00E-08</b>	
AJ565557	AJ565557.1.a.cg.1	NA		9.00E-05	
<b>CU681620</b>	<b>oygd09b07c21r1_m13rev.1.a.cg.1</b>	<b>Superoxide dismutase</b>	<b>2.00E-62</b>	<b>9.50E-05</b>	
<b>FP091069</b>	<b>oyio10b01h21r1_m13r.1.a.cg.1</b>	<b>Dopamine receptor</b>	<b>4.00E-24</b>	<b>9.80E-05</b>	
CX069266	oygd10b09a22r1_m13rev.1.a.cg.1	NA		9.80E-05	
CX069255	CX069255.1.a.cg.1	NA		9.90E-05	

## Figures

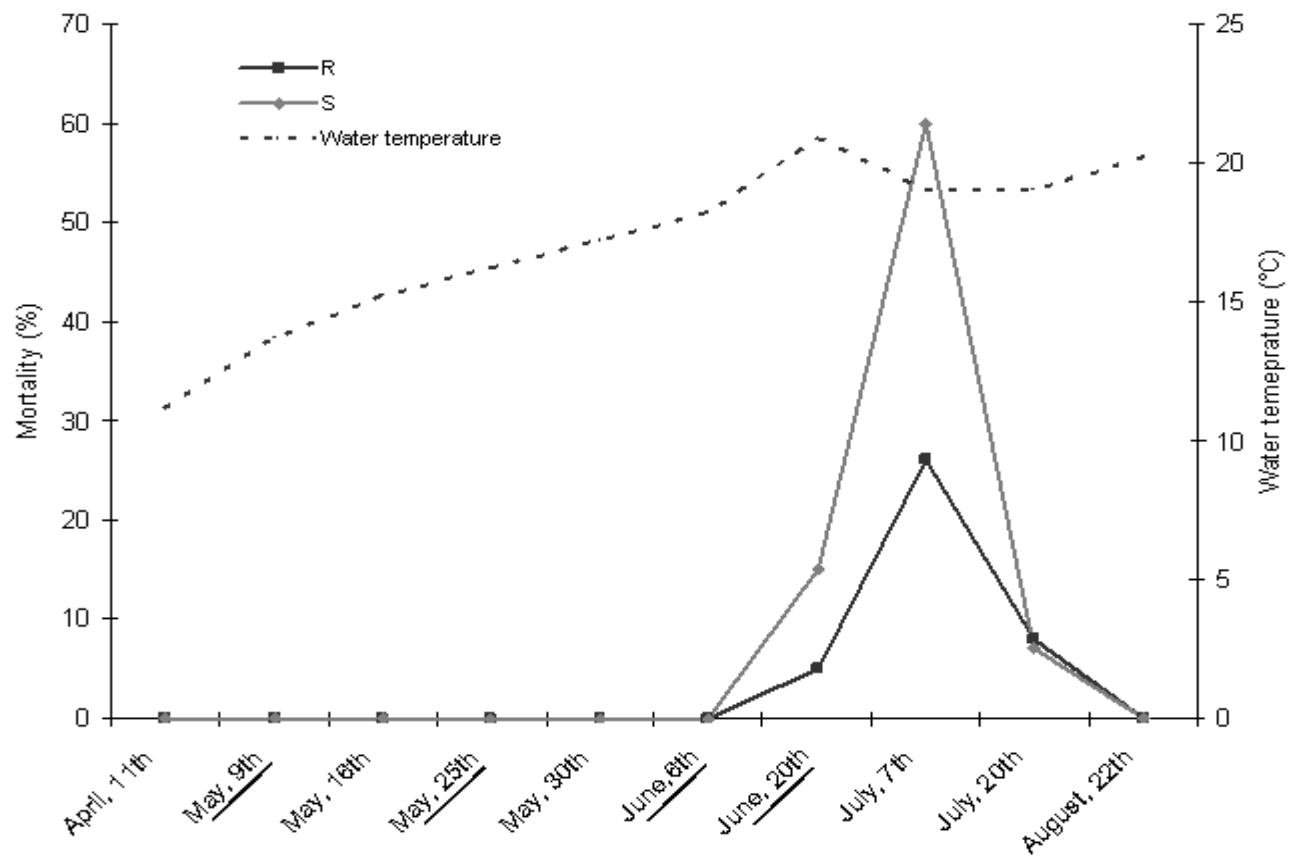


Figure 1: Mortality rates of Resistant (black) and Susceptible (grey) oyster lines and evolution of the water temperature during a field experiment (Fort Espagnol, Brittany, France, 2005). The 4 sampling dates used for microarray analysis are underlined; date 1: May 9, date 2: May 25, date 3: June 6, and date 4: June 20. The other sampling dates were used to assess the mortality rates of the studied oysters, and the water temperature of the field. The box indicates the mortality peak.

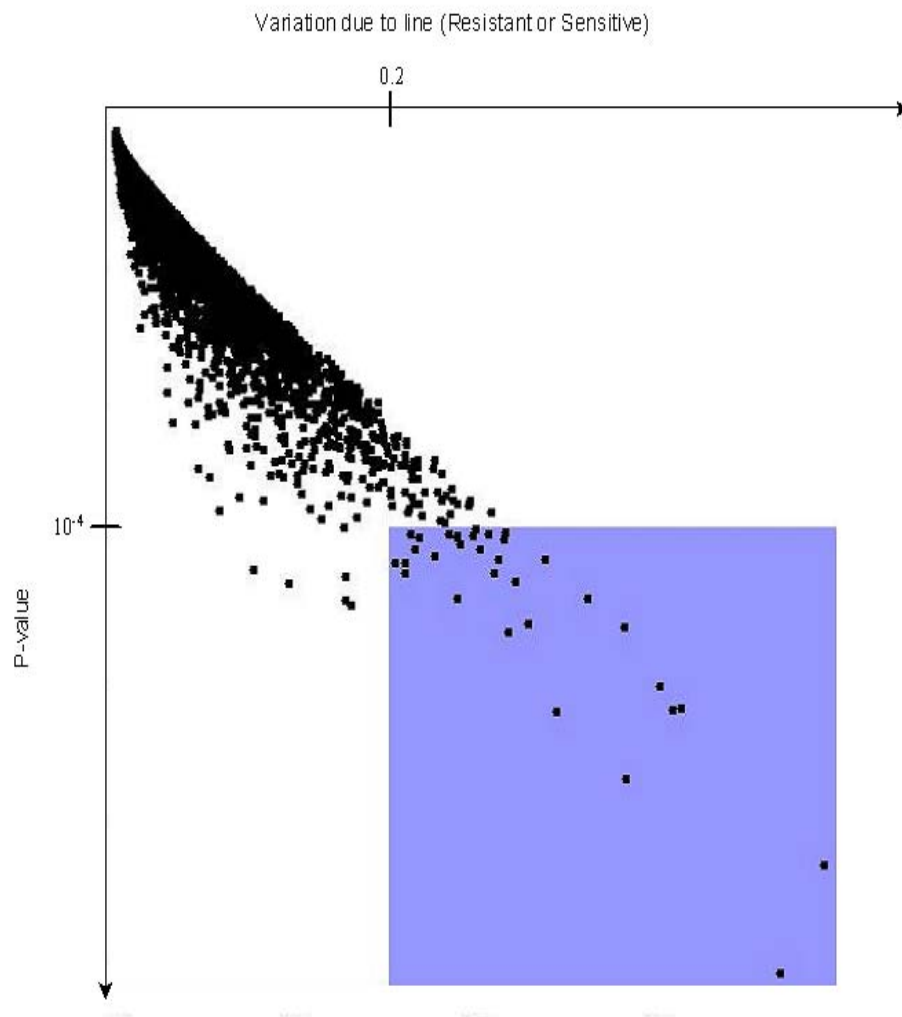


Figure 2: Gene expression analysis of variance with GeneANOVA software (Didier et al., 2002) of the 9058 ESTs studied.

In this two-dimensional visualization, each of the 34 genes is plotted as a point, the abscissa is the variation due to the factor “line” (oyster lines resistant and susceptible to summer mortality) normalized by the total variation of the gene, and the ordinate is the logarithm of the p-value. The shaded area contains the selected genes that had a large proportion of their variation accounted for by the “line” factor, and the lowest p-values (i.e. strongest evidence for differential expression).

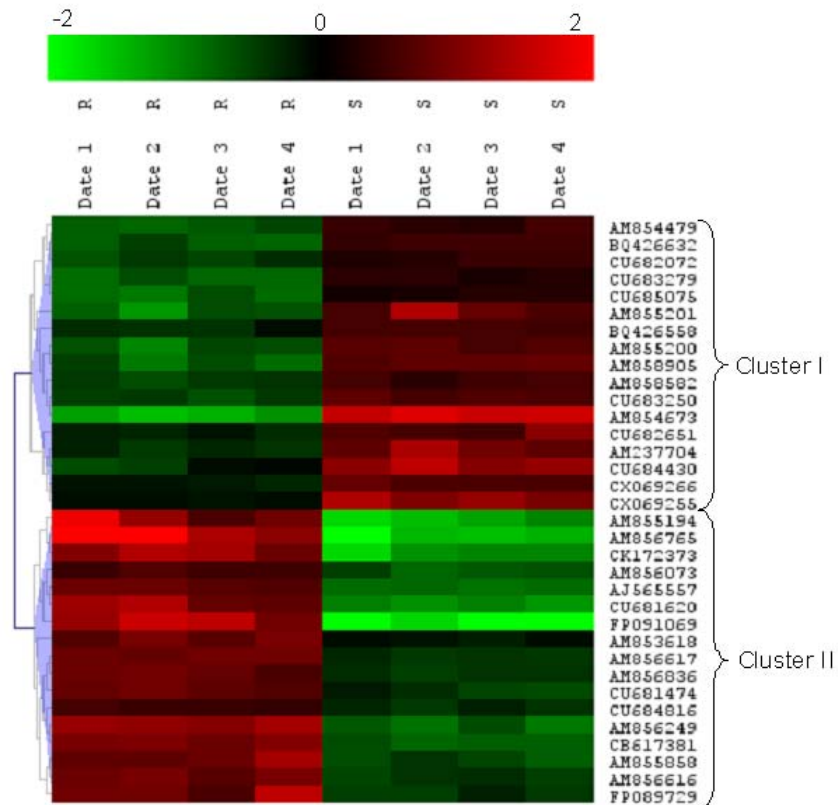


Figure 3: Cluster analysis of expression profiles of the genes selected by GeneANOVA (Didier et al., 2002) for their significant difference of expression between oyster lines resistant and susceptible to summer mortality. Hierarchical clustering by average linkage was carried with Tmev software (<http://www.tm4.org/mev.html>). Columns correspond to the 4 dates of *in situ* sampling for Resistant (R) and Susceptible (S) oyster lines and rows correspond to the selected genes (n=34). Color represents the transformed normalized Cy5 log value obtained for R and S samples, for each date. The variations in transcript abundance are depicted with a color scale, in which shades of red represent increase and shades of green represent decrease in transcript level. Thus, genes under-expressed in R compared to S are found in cluster I, and genes over-expressed in R compared to S are found in cluster II. The dendrogram on the left represents correlation distances between the profiles of studied genes.

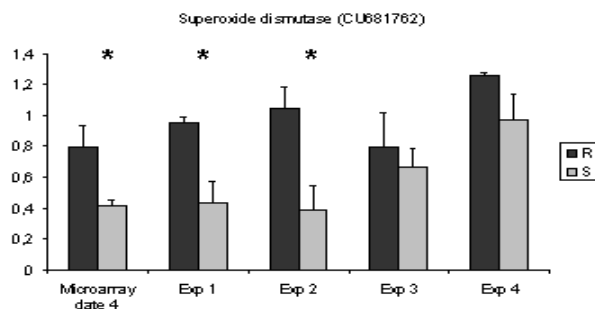
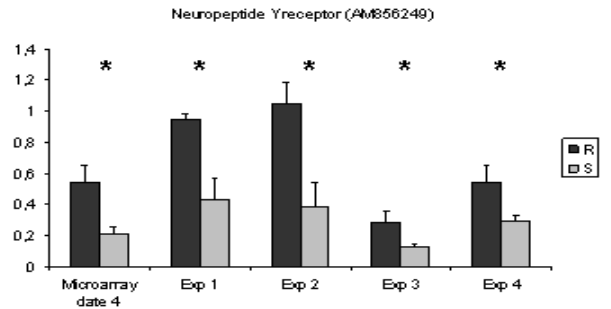
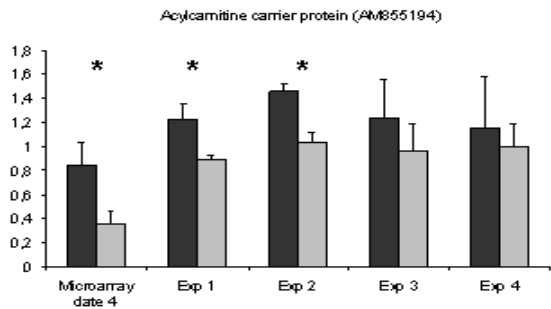
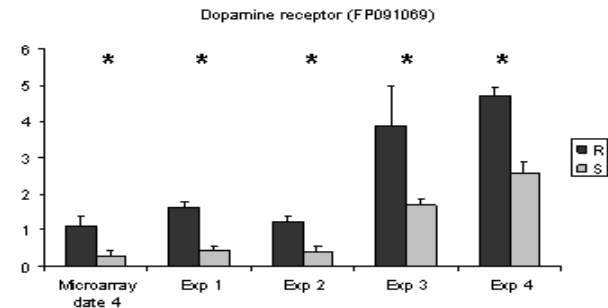
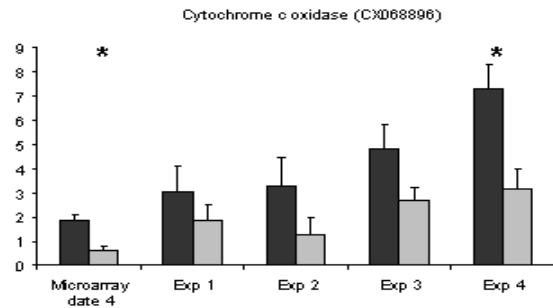
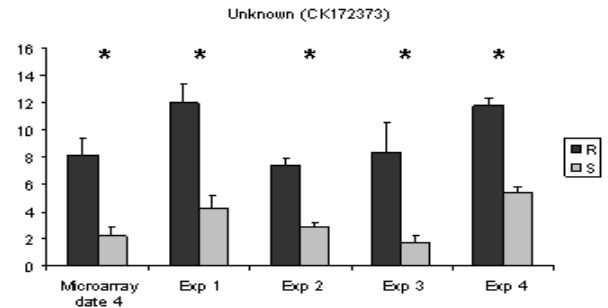
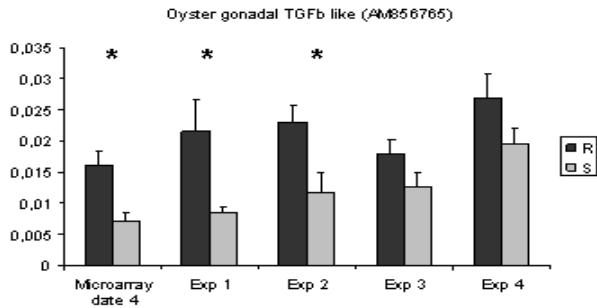
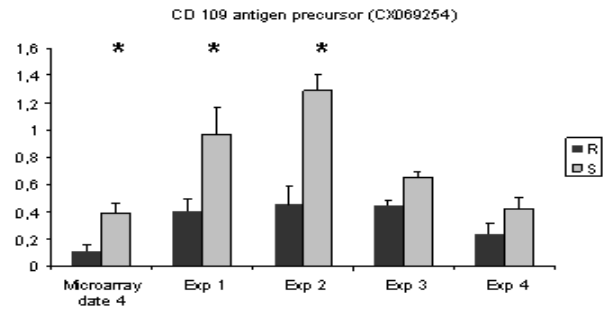
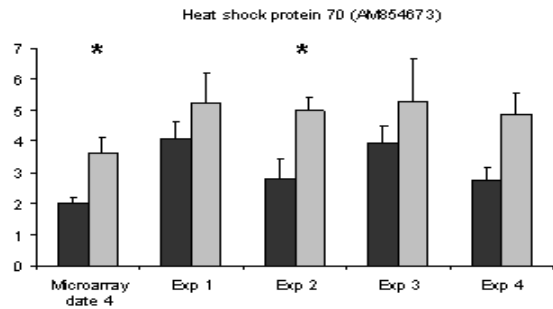


Figure 4: mRNA levels estimated by real-time PCR of 9 selected genes in Resistant (R) and Susceptible (S) oysters relative to Elongation Factor I and to the reference sample used for microarray, sampled from:

- Microarray experiment date 4 (mRNA levels obtained for the 20th June, two weeks before the mortality peak, for R and S lines respectively) (mortality rates: 56% for S, 22% for R).
- Exp 1: spat oysters, 4 months old, second generation, nursery, two weeks before mortality peak (mortality rates: 74% for S, 4% for R).
- Exp 2: spat oysters, 4 months old, third generation, nursery, two weeks before mortality peak (mortality rates: 80% for S, 30% for R).
- Exp 3: adult oysters, 16 months old, second generation, South Brittany field, two weeks before mortality peak (mortality rates: 14% for S, 5% for R).
- Exp 4: adult oysters, 16 month old, second generation, Normandy field, two weeks before mortality peak (mortality rates: 8% for S, 5% for R).

Y axis corresponds to the level of transcripts relative to Elongation Factor I and to the reference sample (also used for microarray experiment). R samples are represented in black and S samples in grey. The asterisk indicates significant differences of mRNA level between R and S at  $p < 0.05$ . Bars represent standard deviation.

## ADDITIONAL FILES

Additional File:

Complete list of the 2114 *Crassostrea gigas* ESTs varying significantly according to the factor "time" (date 1, 2, 3 and 4), independently of the Resistant or Sensitive line.

'Genbank #' corresponds to the Genbank accession number, 'Sigenae Name' corresponds to the ESTs names available in the GigasDatabase ([http://public-contigbrowser.sigenae.org:9090/Crassostrea\\_gigas/index.html](http://public-contigbrowser.sigenae.org:9090/Crassostrea_gigas/index.html)), and 'p-value ANOVA' corresponds to the p-value obtained with the GeneANOVA analysis.