Mass mortality in Pacific oysters is associated with a specific gene expression signature

MAXINE L. CHANEY and ANDREW Y. GRACEY

Marine Environmental Biology, University of Southern California, 3616 Trousdale Pkwy, Los Angeles, CA 90089, USA

Abstract

Mass mortality events occur in natural and cultured communities of bivalve molluscs. The Pacific oyster, *Crassostrea gigas*, is a dominant species in many intertidal locations as well as an important aquacultured bivalve species, and for the last 50 years, adult oysters have suffered frequent and extreme mass mortality events during summer months. To investigate the molecular changes that precede these mortality events, we employed a novel nonlethal sampling approach to collect haemolymph samples from individual oysters during the period that preceded a mortality event. Microarray-based gene expression screening of the collected haemolymph was used to identify a mortality gene expression signature that distinguished oysters that survived the mortality event from those individuals that died during the event. The signature was cross-validated by comparing two separate episodes of mortality. Here, we report that near-mortality oysters can be distinguished from longer-lived oysters by the elevated expression of genes associated with cell death, lysosomal proteolysis, and cellular assembly and organization. These results show the potential utility of nonlethal sampling approaches for investigating the environmental causes of mortality in natural populations in the field, and for predicting when such events could occur and which individuals will be affected.

Keywords: autophagy, death, gene expression, lysosome, microarray, oyster

Received 8 February 2011; revision received 27 April 2011; accepted 4 May 2011

Introduction

Many animals experience mass mortality events in their natural habitat with incidents involving fish (Murray et al. 2003), amphibians (Rachowicz et al. 2006) and marine mammals (Harwood & Hall 1990; Kuiken et al. 1994), often receiving particular attention from the media. Some of these events have significant ecological implications by affecting community structure at various trophic levels (Carpenter 1988; Gomez-Gutierrez et al. 2003) and can have long-term effects (Knowlton 2001). The causes of mass mortality events can sometimes remain cryptic although in many incidences infections by pathogens appear to be the primary cause (Vezzulli et al. 2010). The causes of mass mortality in bivalve populations are of interest because bivalves are often used as sentinels of ecosystem quality, and because mass mortality events can impact the species communities associated with them and the fisheries that are dependent on them. Mass mortality in bivalve communities occurs with varying severity and temporal irregularity. In the mussel, *Mytilus edulis*, mortality rates in excess of 80% were reported in 1981 in Japan and coincided with an unusually hot summer (Tsuchiya 1983). In the Yellow clam, *Mesodesma mactroides*, populations were reduced by at least 80% in its native range in South America after mass mortalities occurring over 3 years in the 1990s during reproductive months (Fiori & Cazzaniga 1999). Viral infections were blamed for mass mortalities of cultured pearl oysters, which exceeded 50% in Japan in 1996 and 1997 (Miyazaki et al. 1999). Pathogen infection and environmental stress have been noted to cause atrophic lesions of the digestive epithelia in Manila clams, *Tapes philippinarum*, which may weaken the animal and be the underlying
cause of observed mass mortalities in Korea (Lee et al. 2001). In all cases, the causes of mass mortality are often difficult to assign because mortality events often proceed quickly and the animals present few symptoms of decline before death.

Mass mortality events occur frequently in the Pacific oyster, *Crassostrea gigas* (Thunberg), which inhabits sheltered bays along the coasts of the Western United States, Mexico, France, Ireland and Japan (Cheney et al. 2000). Natural rates of mortality in this oyster are typically <10% cumulative per year, whereas 30–70% of an assemblage of oysters may die during a mass mortality event (Soletchnik et al. 2005). The first incidences of mass mortality were reported in Japan, the native habitat of the *C. gigas*, in the 1940s and affected both wild populations and harvested beds of adult oysters with particularly severe mortalities reported following the 1958 El Niño (Imai et al. 1965). A similar pattern has been observed in other parts of the world with mortality impacting both wild oysters that naturally recruit to sites and animals that were reared in hatcheries before being transported into the field for grow-out (Cheney et al. 2000), suggesting that the phenomenon is not associated with aquaculture practices. Instead, the onset of mortality is believed to be linked to environmental factors in the field and as such the phenomenon has been termed ‘summer mortality syndrome’ because it is most prevalent once water temperature reaches 18–20°C in summer (Cheney et al. 2000).

The ecology of mass mortality events is complex. *C. gigas* inhabit sheltered bays that exhibit high summer nutrient levels with mortality predominantly impacting reproductively mature, fast-growing animals (Cheney et al. 2000). The correlation between the onset of mortality events and the presence of warm, nutrient-rich water has led to the hypothesis that the affected oysters may be suffering from a type of metabolic imbalance because of an overload in reproductive maturation under these environmental conditions (Imai et al. 1965; Berthelin et al. 2000). Consistent with this hypothesis is evidence that mortality is exacerbated by conditions of reduced oxygen (Soletchnik et al. 2005), as well as siltation and algal blooms (Soletchnik et al. 2007), which may place an additional stress on the metabolism of the oyster. Similarly, studies have shown that lines of oysters in which gametogenesis is slower and reproductive effort less are more resistant to summer mortality, suggesting that the timing and degree of reproductive investment are important factors in the incidence of mortality (Huvet et al. 2010). Susceptibility to summer mortality is also a heritable trait, and efforts to identify resistant and susceptible genotypes have been successful (Samain et al. 2007; Sauvage et al. 2010). Transplantation of these genotypes between different field sites results in different rates of mortality, reinforcing the idea that complex genotype by environment interactions are important in the field (Degremont et al. 2005). While an infectious agent was suspected initially as a causal factor in summer mortality, studies have shown that neither bacterial nor viral infections appear to be primary factors in the incidence of mortality (Garnier et al. 2007). Combined, the current literature suggests that summer mortality is a consequence of multiple environmental factors that may interact in a nonlinear fashion to bring about mortality, with the ability of a particular oyster to adapt to or tolerate these stressors linked to the metabolic poise of the individual, which in the summer months is heavily invested in gonad maturation. It should be noted that scientists’ efforts to simulate summer mortality in the laboratory have been unsuccessful, suggesting that the complex confluence of environmental factors that can occur in a field setting is the primary cause of mortality (Daniel Cheney, pers. comm.).

Molecular approaches have the potential to elucidate the molecular basis of mortality and to explore how the intertidal field environment affects the physiology of its inhabitants. Gene expression profiling of field assemblages of the mussel, *Mytilus californianus*, revealed that the transcriptome responds rapidly to changes in the environmental conditions driven by the ebb and flow of the tide (Gracey et al. 2008). Temperature differences in particular were implicated as being the predominant factor governing differences in the gene expression patterns of mussels living at different positions on the shore. Studies of the transcriptome of the eastern oyster, *Crassostrea virginica*, reveal a similar picture, showing that transcriptional differences between oysters from different locations can serve as a molecular fingerprint providing insights into the quality of their local environment and the effect it has on the physiology of its inhabitants (Chapman et al. 2009, 2011). Thus, sessile intertidal organisms are not passive to changes in their environment, and analysis of their transcriptome can reveal gene expression signatures that can provide linkages between the environment and outcomes at the level of the health and physiology of the organism.

The development and identification of biomarkers is a critical component in environmental monitoring efforts because detection of biomarkers provides an early warning to scientists that exposure to a particular substance, disease agent or environmental factor has occurred and can help to predict important ecological consequences. In the context of the present study, we sought to identify gene expression biomarkers that could differentiate between oysters that would survive or die during a mortality event. Identification of biomarkers may offer molecular insights into the
pathophysiology of mortality, as well as providing tools to investigate the timing of mortality events and how it relates to the prevailing environmental conditions in the field. They may also provide an early warning of conditions that are deleterious to particular inhabitants of the ecosystem. Biomarker discovery was achieved by taking haemolymph samples nonlethally from individual oysters at repeated intervals throughout the summer to create a temporal series of samples for each individual. Mortality was monitored concurrently, which allowed us to identify a set of individuals from which haemolymph samples were collected in the period that preceded their death. Using a microarray-based gene expression screen, we were able to identify a mortality gene expression signature that distinguished oysters that survived the summer from individuals that suffered mortality. The signature was cross-validated by comparing biomarkers generated across two separate mortality events.

Methods

Sampling

Three hundred siblings were selected from an F2 hybrid oyster family created in 2005 from a single cross of F1 siblings arising from a cross of inbred lines 51 and 35 (51 × 35). The crosses were performed by Taylor Shellfish Farms, and the oysters were grown up in the field at Totten Inlet, in the southwestern part of Puget Sound, Washington State. Adult F2 oysters were individually identified by a unique shellfish tag that was glued to the shell. Tagged oysters were kept in black plastic mesh bags, anchored to the sandy beach at the middle intertidal. To prevent overcrowding, each bag contained 50–70 oysters. Oyster body temperatures in each bag were approximated by the deployment of temperature data-loggers (iButton, Maxim Dallas) embedded in silicon sealant inside an oyster shell of similar size to the experimental oysters (Helmuth et al. 2002) with temperature recorded every 1 h.

One challenge that we faced was how to obtain biological samples from individuals at appropriate times during the progression from a healthy oyster through to death. The samples had to be obtained nonlethally so that the future health of the individual could be monitored. One way that this can be accomplished is by taking biopsy samples from mantle tissue, but the procedure is slow because the valves of the shell can only be pried apart after the oyster has been incubated in a bath containing MgCl2 to relax the adductor muscle. Therefore, we turned to an alternative sampling technique in which a sample of haemolymph was collected from each oyster through a hole drilled through the shell above the adductor muscle. In March 2008, the oysters’ ventral shell were predrilled with a small guide hole through which a needle could be inserted into the adductor muscle, using a hand drill and a diamond engraving tool bit (Dremel).

Oysters were sampled on spring tides, on 27 May, 19 June, 14 July, 30 July, 28 August, and 16 October 2008. Oyster bags were removed from the anchored line, and 100 μL of haemolymph was withdrawn from each oyster using a 25-gauge needle that was inserted through the guide hole into the muscle tissue. The sample was expelled into a 2-mL centrifuge tube and immediately frozen on dry ice. Samples were archived at ~80°C for later processing. The oyster bags were then returned to the same location on the shore. At each sample time point, the oysters were carefully examined and dead animals were noted. Oysters that survived the summer were classified as ‘alive’, and those that died were classified ‘mortality’. Mortality was judged to have occurred during the interval between the time point that the animal was identified as dead and the preceding sampling time point.

C. gigas microarray construction

The C. gigas cDNA microarray was constructed from PCR amplicons derived from cDNAs picked from seven high-quality cDNA libraries prepared from adult gill, adductor muscle, mantle tissues and early larval RNA samples. To ensure a good representation of environmentally regulated genes, the adult tissue mRNA was isolated at two time points following exposures on the animals to heat, cold, hypoxia, hypo-osmotic stress, aerial emergence or oxidative stress. The libraries were normalized and serially subtracted so that cDNAs that had been isolated from previous libraries were physically subtracted during the construction of each new library (Carninci et al. 2000); 5’ and 3’ EST sequences were generated from the clones and the sequences annotated using EST-Ferret, a custom annotation pipeline (Gracey et al. 2004). The EST sequences were clustered with CAP3, and the resulting nonredundant contigs were queried against the public protein databases using BLASTX. Each cDNA was assigned a putative identity based on search results against the SwissProt and RefSeq protein databases, and sequences that yielded a hit with an E-value ≤ 1e-10 were annotated using the name of the gene for which it had the greatest identity. Gene ontology terms were assigned to cDNAs with hits in SwissProt by parsing the SwissProt GOA association file with the SwissProt IDs obtained by the BLASTX search. A putative nonredundant set of 11 904 cDNAs was identified and printed on in-house prepared poly-L-lysine-coated microarray slides.
RNA isolation and microarray hybridization

Total RNA was isolated from haemolymph using Trizol (Invitrogen) according to the manufacturer’s instructions. The total RNA was purified further across glass fibre filler columns (Qiagen) according to the manufacturer’s instructions. Amplified RNA was prepared as previously described (Brandish et al. 2005). Briefly, double-stranded cDNA was prepared by reverse-transcribing the entire haemolymph RNA sample (typically 100–500 ng) in a 20-μl reaction containing 20 pmols of T7-dT15VN primer, 50 mM Tris–HCl (pH 8.3), 13 mM KCl, 3 mM MgCl₂, 50 μg dNTP, 20 U of RNaseOut (Invitrogen) and 100 U of MMLV-reverse transcriptase (Epicentre). The RNA and primer mixture were heat-denatured at 65°C for 10 min, then the remaining components were added and the reaction incubated at 40°C for 2 h, and then stopped by heating to 65°C for 15 min. Next, 60 μL of in vitro transcription reaction mixture, containing 53 mM Tris–HCl (pH 7.5), 13 mM NaCl, 8 mM MgCl₂, 5.3% PEG 8000, 2.6 mM spermidine, 3.33 mM each (ATP, GTP, CTP), 2.5 mM UTP, 0.83 mM amino-allyl UTP (Epicentre), 0.12 U organic pyrophosphatase, 20 U RNaseOut, 500 U T7 RNA polymerase (Epicentre), was added and the reaction incubated at 40°C for 18 h. The resulting aRNA was purified using a Qiagen RNeasy kit, and half the aRNA was labelled with Cy5 and the other half with Cy3. Fluorescently labelled aRNA samples were hybridized to the C. gigas array using an interwoven loop design in which alive and mortality samples were directly compared in competitive hybridizations. The loop of hybridizations yielded a balanced design in which each RNA sample was hybridized to either 2 or 4 arrays with fluor-reversal. This loop hybridization design yields improved statistical inference of microarray data (Kerr & Churchill 2001).

Gene expression data analysis

TIFF images of hybridized arrays were captured with an Agilent scanner (Agilent technologies), and spot intensities quantified with Agilent Feature Extraction software (ver. 9.5.1). Spot median pixel intensities without background correction were collected, spatial intensity trends removed and individual channels normalized using joint lowess transformation. Relative expression of each gene in each hybridization loop was estimated using MAANOVA version 0.98–7 for R using an ANOVA model in which dye and array were treated as fixed and mixed effects, respectively, with phenotype class (alive or mortality) as the tested term. The P values were adjusted for false discovery rate (Benjamini & Hochberg 2001), and a statistical threshold of $P < 0.01$ defined ‘gene sets’ that were statistically significantly differentially expressed between the alive and dead class of animals in each data set. The number of genes that passed this threshold was very high in the July data set, so for this data set a secondary threshold was applied which required that genes exhibit a 1.5-fold difference between classes. This additional threshold was applied because the next step in the analysis uses Gene Set Enrichment Analysis (GSEA) (Subramanian et al. 2005), which performs poorly with gene sets that represent more than 5% of the total number of genes in the data set. Next, we used GSEA to investigate the relative distribution of the June gene sets within the rank-ordered list of genes in the July data set, and vice versa, the distribution of the July gene sets within the rank-ordered list of genes in the June data set. This method has advantages over approaches that rely on detecting overlaps between gene lists using Venn diagrams because such methods are highly sensitive to the arbitrary statistical cut-off that is applied when gene lists are created, which will penalize data sets in which phenotypes are not clearly differentiated, or data sets that comprise fewer samples and have lower statistical power. The significance to which a gene set was enriched in the rank-ordered gene lists was calculated by creating 1000 random gene sets, size matched to the actual gene set, and determining the frequency with which the enrichment score was greater than that of the actual gene set, to yield a $P$ value. For each gene set, we used GSEA’s leading-edge statistic to identify the core subset of genes within the gene set, which accounts for the gene set’s enrichment signal. The leading-edge subsets of genes identified in each gene set were merged to create the gene expression mortality signature.

We used Ingenuity Pathways Analysis (Ingenuity Systems, CA) to investigate whether particular biological processes were enriched in the mortality signature. Putative human orthologs of the oyster genes were identified using BLASTX homology searches between the arrayed oyster 5’ EST sequences and the human protein collection in the SwissProt database. We assigned the oyster genes with a putative human gene identifier and used the lists of arrayed genes and

© 2011 Blackwell Publishing Ltd
mortality signature genes as the reference and query gene lists in the analysis. The lists were first dereplicated so that only one instance of a gene identifier was input into the analysis. A right-tailed Fisher’s exact test was used to calculate a \( P \) value that determined the probability that an enriched biological function in the mortality signature was because of chance alone.

**Results**

**Identification of a mortality-associated gene expression signature**

In this experiment, we used gene expression profiling to identify genes that were correlated with field mortality in Pacific oysters and which might serve as biomarkers of a mortality event. We employed an experimental design that is used in quantitative trait locus (QTL) mapping studies and deployed an \( F_2 \) family of oysters that were derived from an earlier cross between two parent inbred \( C.\ gigas \) lines (lines 51 and 35) that exhibited significantly greater field mortality than their \( 51 \times 35 \) hybrids (Dennis Hedgecock, pers. comm.). This experimental intercross design increased the likelihood that individual oysters that we screened would show quantitative differences in mortality resistance. The oysters were grown at Totten Inlet in Puget Sound, which is a commercial grow-out site for Pacific oysters and other shellfish. This site supports rapid growth but is also a site at which high mortality can occur during some summer growing seasons (Cheney et al. 2000, Jonathon Davis, pers. comm.). Estimated body temperatures of the oysters showed that average body temperature increased in the early summer from May to July and that the oysters experience periodic bouts of significantly elevated body temperature (Fig. 1). These warming events occurred during periods when low tides fell on sunny afternoons, but these bouts did not appear to correlate with pulses of increased mortality.

We employed a nonlethal sampling method that was fast and allowed us to collect haemolymph samples from the adductor muscle sinus of individual oysters repeatedly over the summer. As evidence that this sampling regime does not adversely affect the incidence of mortality in oysters, we undertook a similar field experiment in 2007 using the same cohort of \( F_2 \) oysters. Two groups of oysters were tagged, intermingled and grown at Totten Inlet. One group of 236 oysters were sampled for haemolymph at five time points between June and September, while a control group of 136 oysters were not sampled for haemolymph. Just 20 oysters died in the haemolymph sampled group (9.7% cumulative mortality), and 11 died from the control group (7.9% cumulative mortality), indicating that repeated haemolymph sampling does not have a significant effect on the rate of mortality (\( P = 0.7 \), Fisher’s exact test). Because of the low mortality observed in 2007, we repeated the experiment in 2008 and present these data here.

In this study, we sampled animals at 3-to 4-week intervals, on spring tides, with the goal of obtaining samples at time points that were close to periods of mortality. We commenced sampling haemolymph on 27 May 2008 and followed an \( F_2 \) family of 300 oysters over six time points through 16 October 2008. Overall mortality during the summer was 41% (123/300 oysters died) (Fig. 1). We chose to focus our molecular analysis on time points surrounding the month of July because of the availability of a large number of individuals in the mortality phenotypic class within this sampling period. The first data set, referred to as the June data set, contrasted the expression differences of haemolymph samples drawn on 19 June from 22 animals that were dead by 14 July versus samples from 14 animals that were still alive at the end of the summer. The second data set, referred to as the July data set, contrasted the expression differences of haemolymph samples that were drawn on 30 July from seven oysters that were dead by 28 August versus samples from seven oysters that survived the entire summer.

Microarray-based gene expression profiling was used to compare the patterns of expression between the two phenotype classes. A limitation of using haemolymph
as a biological sample is that the yield of RNA from each sample was very low providing sufficient RNA for just a single attempt to produce amplified RNA. Samples were only included in the data set if they yielded sufficient aRNA for hybridization to at least two microarrays with fluor-reversal. In our experience, ~75% of samples reached this threshold when the entire RNA sample was used as the template for the RNA amplification reaction. We used ANOVA to select sets of genes that were significantly up- or downregulated in oysters that would suffer future mortality by the time of the next sampling event versus animals that remained alive through the remainder of the summer. In the June data set, the upregulated and downregulated gene sets comprised 80 and 138 genes, respectively, and in the July data set, they comprised 104 and 130 genes, respectively (Fig. 2a). We ranked the genes in each data set according to whether they were upregulated or downregulated with respect to future mortality and then used GSEA to detect whether the gene sets derived from one data set were statistically enriched in the rank-ordered list of genes from the other data set (Fig. 2b). We detected a significant enrichment of the June upregulated gene set within the rank-ordered July data set ($P < 0.001$), and vice versa between the July upregulated gene set and the rank-ordered June data set ($P < 0.001$) (Fig. 2c). We found no evidence of enrichment for either downregulated gene set in the corresponding rank-ordered gene list for the other data set.

**Functional interpretation of the mortality expression signature**

A core set of 124 genes were identified as contributing to the enrichment of the upregulated gene sets in both the June and July data set (Table S1, Supporting information). The genes that comprised this mortality signature showed a pattern of upregulation in haemolymph samples taken from oysters that would suffer mortality within ~4 weeks of the date on which the sample was drawn (Fig. 2c,d). Using Ingenuity Pathways Analysis, we found that the mortality signature gene list was enriched for genes involved in cell death ($P = 0.004$), protein synthesis ($P < 0.001$), and cellular assembly and organization ($P < 0.001$) (Table 1).

The cell death category includes two members of the Bcl-2 protein family (B2CL1 and B2CL2) that are located in the mitochondrial outer membrane and regulate cytochrome C and reactive oxygen species release from the mitochondria, key steps in the induction of apoptosis (Zamzami et al. 1998). Another member was CED1, which in *caenorhabditis elegans* serves as a cell surface marker that marks apoptotic or corpse cells for engulfment (Zhou et al. 2001). The signature also included growth arrest and DNA damage-inducible protein GADD45 gamma (GA45G) and A-kinase anchor protein 13 (AKP13), both of which have been implicated in the regulation of the apoptotic cascade (Sheikh et al. 2000; Horbinski & Chu 2005). This category included lipopolysaccharide-binding protein (LBP) that binds to bacterial lipopolysaccharides as part of the immune defence response to infection (Zhang et al. 2010).

Closer inspection of the enriched category of genes associated with protein synthesis indicated that the group was dominated with genes involved in protein degradation. For example, the signature contained three isoforms of cathepsin (CATB, CATK and CATL) and legumain (LGMN), which are all lysosomal proteolytic enzymes. The signature also included two genes associated with the response to wounding: transglutaminase (TGM1) that functions in scaffolding of proteins (Griffin et al. 2002) and matrix metalloproteinase 17 (MMP17) that functions in tissue remodelling during healing (Mirastchijski et al. 2010).

The group of genes associated with cellular assembly and organization included many structural genes associated with protein synthesis indicated that the group was dominated with genes involved in protein degradation. For example, the signature contained three isoforms of cathepsin (CATB, CATK and CATL) and legumain (LGMN), which are all lysosomal proteolytic enzymes. The signature also included two genes associated with the response to wounding: transglutaminase (TGM1) that functions in scaffolding of proteins (Griffin et al. 2002) and matrix metalloproteinase 17 (MMP17) that functions in tissue remodelling during healing (Mirastchijski et al. 2010).

The group of genes associated with cellular assembly and organization included many structural genes such as myosin heavy chain (MYS), myosin essential light chain (MLE), two myosin regulatory light chains (MLRA and MLRC), myophilin (MYPH), paramyosin (MYSP), filamin (FLNA) and tropomyosin (TPM). In addition, three isoforms of troponin (TNNI, TNMM and TNNT) were present in the mortality signature.

Other genes of note that were present in the signature included two key metabolic enzymes: phosphoenolpyruvate carboxykinase (PCKGC) and adenosine kinase (ADK1). The remaining member genes participate in a wide variety of processes and a shared functional theme was not apparent.

**Discussion**

Deteriorating health has system-wide effects on animal performance, impacting all levels of physiology and manifesting in systems beyond the primary lesion. For this reason, accessible biological fluids such as blood are important surrogate systems for monitoring animal health (DePrimo et al. 2003; Tang et al. 2003; Whitney et al. 2003; Rubins et al. 2004; Borovecki et al. 2005; Li et al. 2006). In this study, we searched for biomarkers in haemolymph fluid, and our findings reveal that specific gene expression changes can be detected in haemolymph drawn from oysters that would die within 4 weeks of the date on which the haemolymph sample was taken. Oyster haemolymph contains a population of hemocytes that consist of at least three cell types and are believed to have a primary phagocytic role in host defence (Anderson et al. 1999). Hemocytes have been used previously as a proxy to investigate oyster health.
because they show characteristic responses to a variety of environmental parameters including heavy metal exposure (Gagnaire et al. 2004), oxidative stress (Gielazyn et al. 2003), salinity and temperature changes (Corporeau & Auffret 2003) and mechanical disturbance (Lacoste et al. 2002). Furthermore, changes in hemocyte cell state have also been correlated with field mortality, with hemocytes exhibiting diminished expression of oxidative stress response proteins and evidence of elevated DNA damage (Uchimura et al. 2003). Given their primary role is in defence, hemocytes seem a candidate cell type to respond to systemic changes in health.

Fig. 2 Identification of a common gene expression signature associated with mortality in two data sets. Box whisker plots of statistically significant gene sets (P < 0.01) that were induced in the (a) June, and (b) July data sets. (c) Data mining strategy based on GSEA. The genes in each data set were rank-ordered according to correlation with mortality (P-value), and then GSEA was used to assess the extent to which the June mortality gene set was enriched in the July data set, and the July mortality gene set was enriched in the June data set. The vertical lines indicate the rank position of each gene in the opposing data set. Integration of the two data sets revealed a subset of 124 genes that were significantly associated with mortality in both expression data sets. Their rank positions in each data set are indicated as horizontal lines. (d) Heatmaps show the relative expression of the 124 gene signature (rows) across the 36 and 14 animals (columns) that comprised the June and July data sets, respectively. The vertical white line delineates the alive (left) and mortality (right) class of individual samples in each data set. Red or green colour indicates higher or lower levels of expression relative to the mean expression of that gene in each data set. The standard deviation from mean, σ, is indicated.
We do not have the ability to predict when an oyster mortality event will occur. Studies indicate that genotype plays a role in determining an individual oyster’s sensitivity to mortality (Sauvage et al. 2010), but environment also has a significant effect on when or whether mortality occurs (Cheney et al. 2000).

Table 1  Genes that are members of functional groups that are enriched in the gene expression mortality signature

<table>
<thead>
<tr>
<th>Functional group</th>
<th>Gene name</th>
<th>Gene symbol</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A-kinase anchor protein 13</td>
<td>AKP13</td>
<td>HS169501, HS169502</td>
</tr>
<tr>
<td></td>
<td>Bcl-2-like protein 1</td>
<td>B2CL1</td>
<td>HS130715, HS130716</td>
</tr>
<tr>
<td></td>
<td>Bcl-2-like protein 2</td>
<td>B2CL2</td>
<td>HS128953, HS128954</td>
</tr>
<tr>
<td></td>
<td>CCAAT/enhancer-binding protein gamma</td>
<td>CEBPAG</td>
<td>HS155435,</td>
</tr>
<tr>
<td></td>
<td>Cell death abnormality protein 1</td>
<td>CED1</td>
<td>HS212871, HS212872</td>
</tr>
<tr>
<td></td>
<td>Growth arrest and DNA damage-inducible protein</td>
<td>GA45G</td>
<td>HS177993, HS177994</td>
</tr>
<tr>
<td></td>
<td>Lipopolysaccharide-binding protein</td>
<td>LBP</td>
<td>HS166583, HS166584</td>
</tr>
<tr>
<td></td>
<td>Actophorin</td>
<td>ACTP</td>
<td>HS199639, HS199640</td>
</tr>
<tr>
<td></td>
<td>Calponin-3</td>
<td>CN3N</td>
<td>HS131821, HS131822</td>
</tr>
<tr>
<td></td>
<td>Glutamate receptor 1</td>
<td>GRIA1</td>
<td>HS170907, HS170908</td>
</tr>
<tr>
<td></td>
<td>Kyphoscoliosis peptidase</td>
<td>KY</td>
<td>HS235972, HS235973</td>
</tr>
<tr>
<td></td>
<td>Muscle LIM protein Mlp84B</td>
<td>MLP2</td>
<td>HS116932</td>
</tr>
<tr>
<td></td>
<td>Muscle M-line assembly protein unc-89</td>
<td>UNC89</td>
<td>HS216967, HS216968</td>
</tr>
<tr>
<td></td>
<td>Myophilin</td>
<td>MYPH</td>
<td>HS186735, HS186736</td>
</tr>
<tr>
<td></td>
<td>Myosin essential light chain, striated adductor muscle</td>
<td>MLE</td>
<td>HS120164, HS120165</td>
</tr>
<tr>
<td></td>
<td>Myosin heavy chain, striated muscle</td>
<td>MYS</td>
<td>HS180121</td>
</tr>
<tr>
<td></td>
<td>Myosin regulatory light chain A, smooth adductor muscle</td>
<td>MLRA</td>
<td>HS127636, HS127637</td>
</tr>
<tr>
<td></td>
<td>Myosin regulatory light chain, striated adductor muscle</td>
<td>MLRC</td>
<td>HS243255</td>
</tr>
<tr>
<td></td>
<td>Paramyosin</td>
<td>MYSP</td>
<td>HS128732, HS128733</td>
</tr>
<tr>
<td></td>
<td>Protein odd-skipped-related 1</td>
<td>OSR1</td>
<td>HS215585, HS215586</td>
</tr>
<tr>
<td></td>
<td>Rootletin</td>
<td>CROCC</td>
<td>HS172725</td>
</tr>
<tr>
<td></td>
<td>Sarcomplasmic calcium-binding protein</td>
<td>SCP</td>
<td>HS201916, HS201917</td>
</tr>
<tr>
<td></td>
<td>Synaptotropin-2</td>
<td>SYN2P</td>
<td>HS176329, HS176330</td>
</tr>
<tr>
<td></td>
<td>Tropomyosin</td>
<td>TPM</td>
<td>HS163450, HS163451</td>
</tr>
<tr>
<td></td>
<td>Troponin C</td>
<td>TNNC</td>
<td>HS134508, HS134509</td>
</tr>
<tr>
<td></td>
<td>Troponin 1</td>
<td>TNN1</td>
<td>HS144383, HS144384</td>
</tr>
<tr>
<td></td>
<td>Troponin T, skeletal muscle</td>
<td>TNNT</td>
<td>HS209964, HS209965</td>
</tr>
<tr>
<td></td>
<td>Tubulin alpha-2/alpha-4 chain</td>
<td>TBA2</td>
<td>HS125492, HS125493</td>
</tr>
<tr>
<td></td>
<td>Filamin-A</td>
<td>FLNA</td>
<td>HS206344, HS206345</td>
</tr>
<tr>
<td>Protein synthesis</td>
<td>ATP-binding cassette subfamily G member 1</td>
<td>ABCG1</td>
<td>HS126413</td>
</tr>
<tr>
<td></td>
<td>Cathepsin B</td>
<td>CATB</td>
<td>HS141294, HS141295</td>
</tr>
<tr>
<td></td>
<td>Cathepsin K</td>
<td>CATL</td>
<td>HS160954, HS160955</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L</td>
<td>CATL</td>
<td>HS129715, HS129716</td>
</tr>
<tr>
<td></td>
<td>Cholecystokinin receptor</td>
<td>CCKAR</td>
<td>HS172981, HS172982</td>
</tr>
<tr>
<td></td>
<td>Deoxyhypusynase synthase</td>
<td>DHYS</td>
<td>HS209302</td>
</tr>
<tr>
<td></td>
<td>Eukaryotic translation initiation factor 4 gamma 1</td>
<td>IF4G1</td>
<td>HS231161, HS231162</td>
</tr>
<tr>
<td></td>
<td>Glutamine synthetase 2 cytoplasmic</td>
<td>GLNA2</td>
<td>HS111068, HS111069</td>
</tr>
<tr>
<td></td>
<td>Legumain</td>
<td>LGMN</td>
<td>HS158882, HS158883</td>
</tr>
<tr>
<td></td>
<td>Matrix metalloproteinase-17</td>
<td>MMP17</td>
<td>HS191831, HS191832</td>
</tr>
<tr>
<td></td>
<td>Neuroreptide S receptor</td>
<td>NPSR1</td>
<td>HS123632, HS123633</td>
</tr>
<tr>
<td></td>
<td>Protein-glutamine gamma-glutamyltransferase K</td>
<td>TGM1</td>
<td>HS212804, HS212805</td>
</tr>
<tr>
<td></td>
<td>Rho-related protein racG</td>
<td>RACG</td>
<td>HS172425, HS172426</td>
</tr>
<tr>
<td></td>
<td>Translation initiation factor elf-2B subunit gamma</td>
<td>E12BG</td>
<td>HS144339</td>
</tr>
<tr>
<td></td>
<td>Tryptophanyl-rRNA synthetase, cytoplasmic</td>
<td>SYWC</td>
<td>HS147659</td>
</tr>
<tr>
<td></td>
<td>Von Hippel-Lindau disease tumour suppressor</td>
<td>VHL</td>
<td>HS131322</td>
</tr>
<tr>
<td></td>
<td>Zinc metalloproteinase nas-13</td>
<td>NAS13</td>
<td>HS125375, HS125376</td>
</tr>
<tr>
<td>Other</td>
<td>Phosphoenolpyruvate carboxykinase, cytosolic [GTP]</td>
<td>PCKGC</td>
<td>HS195189, HS195190</td>
</tr>
<tr>
<td></td>
<td>Adenosine kinase 1</td>
<td>ADK1</td>
<td>HS230416, HS230417</td>
</tr>
</tbody>
</table>

The complete list of genes in the mortality signature is available in Table S1 (Supporting information).
Environmental differences between summers may explain why the cohort of oysters we were monitoring suffered only 9.7% mortality in 2007 but 41% mortality in summer 2008, although we cannot exclude the possibility that the increased size and age of the oysters by 2008 was a contributing factor. This variation in mortality incidence means that samples have to be taken at a number of time points with the intent that some collected samples will coincide with a mortality event. In this study, we sampled oysters at 3- to 4-week intervals because access to the animals was constrained by tides and their remote location in the field. As a result, we were unable to assign a specific time of death to each oyster. Despite the low resolution of time points, we were able to detect a gene expression signature that distinguished animals that died during this 4-week window from those that survived throughout the summer. This suggests that deterioration of the health of the oyster and associated changes in gene expression commenced weeks rather than days or hours before mortality occurs. This means that the environmental factors that led to the onset of mortality and the mortality event itself may be temporally separated in time, which presents a challenge for understanding the specific factors that induced mortality. However, the fact that mortality is not sudden also raises the possibility that early detection of mortality biomarkers could allow manipulative experiments to be undertaken to investigate whether the rate of mortality can be ameliorated, for example by transplanting animals to an alternative site, or if the animal's fate is fixed and cannot be reversed.

One shortcoming of our study design was that we were unable to assign an accurate date of mortality to each individual. This probably degraded our statistical analysis because animals that died at any time between sampling intervals were grouped together in the mortality class of individuals, and it is reasonable to assume that the expression signature of animals that were closer to death would be different from those that were sampled at a time point that was further from death. Despite this shortcoming, a signature that was common to oysters that suffered mortality was detected. However, greater resolution regarding the time of mortality might have provided a time-resolved profile of the molecular changes that are associated with the progression towards death and provided more insights into the pathophysiology of premature mortality. Further studies are required in which animals are sampled at much greater frequency, thus shrinking the time of the window during which they died, or using an experimental design in which the health of individual animals is scored on a daily basis. The latter option would be difficult to implement in this low intertidal field location and would be more appropriate for a laboratory setting.

An alternative study design would be to analyse the expression profile of individual oysters early in the season under benign field conditions and then to interpret these data in the context of ultimate fate of the individual during the ensuing summer season. This design is easier to implement because haemolymph samples have to be collected only once; however, a trade-off is that some of the expression differences that differentiate surviving from mortality individuals may only be evident under field conditions that induce summer mortality.

The mortality gene expression signature was identified by integrating two temporally separated data sets to uncover a signature that is representative of more than one mortality incident. The mortality signature was comprised entirely of genes whose expression is elevated in animals that would die, and we found no evidence for a common pattern of suppressed gene expression. This suggests that the onset of mortality has a stronger effect on the induction of genes and that mortality is not associated with a paucity of expression of specific genes. The discovery of genes associated with cell death in the mortality signature validated our sampling strategy and suggests that mortality of the whole organism is preceded by the activation of death processes in individual cells. The signature was enriched for lysosomal proteases, including three different cathepsin genes, and there is strong evidence for a role for cathepsins in the apoptosis signalling cascade (Chwieralski et al. 2006), and with activation of the lysosomal pathway preceding the apoptotic cascade (Joy et al. 2010). Lysosomes are responsible for the breakdown not only of cellular waste products but also of material that is brought into the cell by endocytosis. This is relevant because the most abundant type of cell in oyster haemolymph is granulocyte that uses pseudopodia to engulf damaged cells, debris and bacteria (Donaghy et al. 2009). The mortality signature included a suite of cell structural genes, and we can speculate that this might be linked to increase in the relative abundance or activity of granulocytes because protrusion of pseudopodia requires extensive regulation of the cytoskeleton (Weaver 2005). Supporting this hypothesis, the signature included CED-1, a receptor that is expressed by engulfing cells and which recognizes cell corpses for engulfment (Reddien & Horvitz 2004). Therefore, one interpretation of the composition of the mortality signature is that autophagy, the lysosomal degradation pathway by which organisms remove damaged cells and cellular components (Levine & Kroemer 2008), is activated and may serve as a marker for the onset of mortality. However, it remains unclear whether the hemocytes are degrading their own cell components as they follow a trajectory towards death, or whether autophagy is acting on debris arising from damaged tissues that were engulfed, or a combination of both processes.
Because we screened an $F_2$ hybrid oyster family, the gene expression changes that we observed may reflect differences in the underlying genetic architecture of mortality-susceptible versus mortality-resistant individuals. However, the presence of genes associated with cell death and autophagy would suggest that at least some proportion of genes in the signature is symptomatic of oysters that are near mortality. Furthermore, the types of genes that make up the mortality signature are different to those that have been identified previously in gene expression screens of gonad from resistant versus susceptible lines of oysters in France (Huvet et al. 2004; Fleury et al. 2010). This difference may be explained by the fact that genotype-associated expression traits are not always shared across tissues (Petretto et al. 2008). For example, the molecular changes associated with the onset of coral bleaching have been investigated by routinely taking samples of coral fragments from multiple colonies throughout the year and then interpreting their gene expression profiles with respect to occurrence of bleaching events in individual colonies (Seneca et al. 2010). The major challenge with the implementation of these approaches is the nonlethal collection of tissue samples from organisms at appropriate times prior to the presentation of the trait that is under investigation.

The hemocyte gene expression pattern that we identified serves as a biomarker of deteriorating health, but our data did not offer significant insights into what causes mortality or the source of the primary lesion. Instead, clues to the primary cause of mortality probably lie in a tissue or cell type other than haemolymph. Studies of other tissues will inevitably require lethal-sampling of the animal, but such efforts could be aided by the use of haemolymph profiling, because detection of the mortality gene expression signature in the haemolymph of a lethally sampled oyster could be used as evidence that the individual was near mortality and therefore an appropriate candidate for further investigations. Such data would provide greater insights into the pathophysiology of mortality by revealing whether shared or specific gene expression programs are activated across tissues in the time leading up to mortality. Ideally, this would be implemented in a laboratory setting, although to date efforts to simulate a mortality event have not been successful. Regardless, identification of the causes of sudden mortality events will require detailed integration of molecular data with time-resolved environmental and mortality data.

Acknowledgements

We sincerely acknowledge the logistical support provided by Jonathon Davis and Taylor Shellfish Farms without whom this work would not have been possible. We also thank Dennis...
Hedgecock and Daniel Cheney for providing the F2 hybrid oysters and for their insights into genotype and environmental interactions on mortality. This research was supported by National Oceanic and Atmospheric Administration National Sea Grant award # NAO07OAR4170435 to AYG.

References


Donaghy L, Kim BK, Hong HK, Park HS, Choi KS (2009) Flow cytometry studies on the populations and immune parameters of the hemocytes of the Suminoe oyster, Crassostrea ariakensis. Fish Shellfish Immunology, 27, 296–301.


© 2011 Blackwell Publishing Ltd
The primary focus of A.G.'s research is to understand the effect of factors such as temperature, oxygen concentration, and salinity on the physiological processes of animals, and how these organism-environment interactions affect performance, survival and distribution. His studies explore the nature of adaptive responses and the evolutionary development of stress tolerance in a range of species spanning multiple phyla, including freshwater and marine fish, nematode worms, bivalve mollusks and hibernating mammals. Since phenotype results ultimately from the combined expression of genes and gene complexes, understanding patterns of gene expression evoked during environmental change yields insights regarding the mechanisms underlying adaptation from the cellular to the whole organism level.

Data accessibility

The accession numbers associated with the mortality signature member genes are provided in the Table S1 (Supporting information). The microarray data have been deposited in the ArrayExpress public database under accession number E-MTAB-548.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Complete list of transcripts that comprise the gene expression mortality signature.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.