Polymorphism in a serine protease inhibitor gene and its association with disease resistance in the eastern oyster (Crassostrea virginica Gmelin)

Haiyang Yu a,b, Yan He a,b, Xiaoxue Wang b, Quanqi Zhang a, Zhenmin Bao a, Ximing Guo b,*

a College of Marine Life, Ocean University of China, Qingdao, Shandong 266003, China
b Haskin Shellfish Research Laboratory, Institute of Marine and Coastal Sciences, Rutgers University, Port Norris, NJ 08345, USA

A R T I C L E   I N F O
Article history:
Received 23 November 2010
Accepted 20 December 2010
Available online 6 January 2011

Keywords:
Serine protease inhibitor
Perkinsus marinus
Disease resistance
Oyster
Crassostrea virginica

A B S T R A C T
Serine protease inhibitors (SPIs) are a superfamily of structurally related but functionally diverse proteins found in almost all organisms ranging from viruses to humans. Some of them play important roles in host defense. A recently identified SPI from the eastern oyster (Crassostrea virginica), cvSI-1, has been shown to inhibit the proliferation of the Dermo pathogen Perkinsus marinus in vitro, although direct evidence linking it to disease resistance is lacking. In this study, we identified polymorphism in the cvSI-1 gene and studied its association with improved survival after disease-caused mortalities and in disease-resistant eastern oyster strains. Full-cDNA sequence of cvSI-1 was sequenced in a diverse panel of oysters, revealing 12 single-nucleotide polymorphisms (SNPs) in the 273 bp coding region: five were synonymous and seven non-synonymous. The Dn/Ds ratio, 1.4, suggests that cvSI-1 is under positive selection. Selected SNPs were genotyped in families before and after disease-caused mortalities as well as in disease-resistant and susceptible strains. At SNP198, the C allele consistently increased in frequency after mortalities that are caused primarily by Dermo and possibly also by MSX. Its frequency in the disease-resistant strain is significantly higher than that in the susceptible strains and the base population from which the selected strains were derived. These results indicate that polymorphism at cvSI-1 is associated with Dermo (possibly also MSX) resistance in the eastern oyster. SNP198 is a synonymous mutation, and its association with disease resistance may be due to its close linkage to a functional polymorphism nearby.

© 2010 Elsevier Ltd. All rights reserved.

1. Background

The eastern oyster (Crassostrea virginica Gmelin), which supports important fishery and aquaculture industries in the United States, is seriously affected by two major diseases: MSX (caused by the protozoan Haplosporidium nelsoni) and Dermo (caused by the protozoan Perkinsus marinus) [1]. Each of the two diseases may kill up to 90% of the naive oysters. These two diseases are considered to be one of the leading causes for the collapse of the oyster populations and fisheries in the mid-Atlantic region [2]. The identification of oyster genes that are involved in disease resistance should improve our understanding of host defense in oysters and contribute to our efforts to develop disease-resistant stocks.

Serine protease inhibitors (SPIs) are a superfamily of structurally related proteins that are found in almost all organisms ranging from viruses to humans [3,4]. They have diverse functions and are involved in many biological pathways and processes. As suggested by their name, most SPIs inhibit the proteolytic activity of serine proteases. Any process that involves serine proteases, including digestion, blood clotting, inflammation and immune responses, may be regulated by SPIs [5]. Some SPIs such as HSP47 may not have inhibitory effects on proteases but function as storage and chaperone proteins [6]. In immune response, SPIs may regulate the Toll pathway [5] or directly inhibit serine proteases from invading pathogens that are often key virulence or pathogenic factors [7,8]. Thus, some SPIs may be important for disease resistance in oysters.

A serine protease inhibitor, cvSI-1, has been identified and characterized in the eastern oyster [9,10]. It has strong inhibitory effects on serine proteases of the oyster pathogen P. marinus and the proliferation of the pathogen itself in vitro. CvSI-1 is expressed at higher levels in oysters selected for Dermo resistance than in unselected controls, although the selected oysters were from Louisiana and unselected oysters were from Maine, and there was no change in cvSI-1 expression level in either population upon P. marinus challenge [10]. Thus, while it is possible or even likely that cvSI-1 is involved in host defense against Dermo, direct evidence...
linking it to improved survival or Dermo resistance is lacking. To further understand if cvSI-1 is indeed involved in disease resistance at organism level, we studied polymorphism in the cvSI-1 gene and analyzed its possible association with improved survival or disease resistance in the eastern oyster. We sequenced the full coding sequence of cvSI-1 in a diverse panel of oysters and identified 12 single-nucleotide polymorphisms (SNPs). Selected SNPs were genotyped in families before and after disease-caused mortality, as well as in disease-resistant and susceptible strains derived from the same base population. Here we present evidence that variation in cvSI-1 is associated with disease resistance in the eastern oyster.

2. Materials and methods

2.1. Oysters, families and populations

Thirty oysters were used for resequencing cvSI-1 and SNP discovery. The oysters were taken at random from three different geographic populations: Mobile Bay (Alabama), Delaware Bay (New Jersey) and Martha’s Vineyard (Massachusetts), ten from each population. To determine if variation at cvSI-1 is associated with improved survival under diseases, selected SNPs were genotyped in three families before and after disease-caused mortalities. The three families were F2 (HB2 and XG3) or backcross (BCM4) families created from selected disease-resistant and susceptible strains derived from the wild base population from which these strains were derived. Oysters from the three families were deployed at Cape Shore, NJ, where Dermo is enzootic. Mortalities were monitored and samples were regularly collected. Before and after mortality samples were identified on mortality curves and used for this study. They bracketed 50–80% mortalities that were mainly caused by Dermo.

To further demonstrate that polymorphism at cvSI-1 is associated with disease resistance, SNPs were genotyped in disease-resistant (NEH) and susceptible (FMF and UMFS) strains as well as the wild base population from which these strains were derived. NEH was a disease-resistant strain selected at Rutgers University since 1960 that had shown strong resistance to MSX [11] and moderate resistance to Dermo [12]. FMF and UMFS were selected strains from Frank M. Flower Oyster Company (NY) and University of Maine, respectively. They had been selected for superior growth and resistance to JOD (juvenile oyster disease, caused by Roseovarius crassostreae), but both were susceptible to Dermo and MSX.

2.2. Resequencing and SNP discovery

For resequencing, total RNA was extracted from gills or adductor muscles of 30 oysters from the panel using the Promega SV Total RNA Isolation Kit. cDNA was synthesized with the Qiagen Quanti-Tect Reverse Transcription Kit. Resequencing primers were designed using Primer Premier 5.0 based on cvSI-1 sequence (gi|71796854|gb|DQ902546.1) in GenBank covering the entire coding sequence (Table 1). PCR was performed in a 20 μl reaction mixture consisted of 100 ng cDNA, 2 μl 10 × PCR Buffer (20 mM Tris–HCl pH 8.3, 500 mM KCl, 15 mM MgCl2), 0.4 μl dNTP (2.5 mmol/L), 0.8 μl of each primer (5 μmol/L), 1 μl bovine serum albumin (0.5 g/L) and 1 U GoTaq DNA polymerase (all from Promega). Reaction conditions for PCR were as follows: 95 °C for 5 min; 35 cycles of 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 1 min; and a final extension at 72 °C for 10 min. PCR products were purified and directly sequenced with ABI 3730 sequencers at the High-Throughput Genomics Unit (HTGU), University of Washington. Known positive and negative controls were included for sequencing, and all samples were sequenced in both directions.

2.3. SNP assay design, genotyping and analysis

For genotyping, genomic DNA was extracted from adductor muscles of each oyster with the Omega Biotech D3373-02 E.Z.N.A. Mollusc DNA Kit, quantified on a nanodrop spectrophotometer and verified for product integrity on agarose gels. Two SNPs (Table 1) were selected for genotyping and characterization with the high resolution melting (HRM) assay using Syto 9 green fluorescent stain. The HRM assay is based on the amplification of a short DNA fragment where different alleles of an SNP give different melting profiles detectable on a real-time PCR system. HRM primers were designed using Primer Premier 5.0 following the CorProtocol for HRM analysis [13]. Product size was set between 60 and 100 bases. Primer size is set to be between 18 and 27 bases long, with melting temperatures between 57 and 63 °C, and GC content between 20% and 80%. Default values were used for other parameters such as max 3’ stability and max mis-priming. Secondary structure of the amplicon was analyzed using the DINAmpel Server (http://www.bioinfo.rpi.edu/applications/hybrid/two-state-fold.php) [14].

PCR mixture consisted of 10 ng of genomic DNA, 200 nM of each primer, AmpliTaq Gold 360 Master Mix (containing 0.5 μM Syto 9, Applied Biosystems Inc.) and PCR grade water in a volume of 20 μl. PCR cycling and HRM analysis were performed on an Applied Biosystems 7500 Fast real-time PCR system (Applied Biosystems Inc.). PCR amplification was conducted with the following protocol: 95 °C for 5 min; 45 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min; and final incubation at 72 °C for 10 min. For melting curve analysis, PCR products were heated from 65 °C to 95 °C, at the rate of 1 °C per second. Normalized melting curve, temperature-shift curve and difference plots were generated with the High Resolution Melting Analysis 2.0 software provided with the ABI 7500 Fast real-time PCR system. Genotypes were automatically called or manually assigned when necessary. Known genotypes identified by sequencing were included as positive controls. The software GENEPOP on the web (http://genepop.curtin.edu.au/) was used to identify deviations from Hardy–Weinberg equilibrium. Frequency difference between populations was examined by chi-square test.

3. Results

3.1. Sequences and SNPs

Of the 30 oysters used for resequencing, 12 oysters failed in PCR amplification or sequencing reaction, and 18 oysters produced high quality sequences covering all three geographic populations. Analysis of the 18 sequences revealed 12 SNPs in the 273 bp coding region, corresponding to about one SNP per 23 bp (Fig. 1). Minor allele frequency ranged from 0.13 to 0.42 (Table 2). Of the 12 SNPs, five were synonymous and seven were non-synonymous (Table 2).
Table 2

<table>
<thead>
<tr>
<th>SNP</th>
<th>Codon</th>
<th>Amino acid</th>
<th>Minor allele frequency</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP94</td>
<td>TCT/TTC</td>
<td>Phe</td>
<td>0.38</td>
<td>S</td>
</tr>
<tr>
<td>SNP51</td>
<td>AGC/ACA</td>
<td>Thr</td>
<td>0.30</td>
<td>S</td>
</tr>
<tr>
<td>SNP52</td>
<td>GTT/TTT</td>
<td>Val/Phe</td>
<td>0.30</td>
<td>NS</td>
</tr>
<tr>
<td>SNP94</td>
<td>TCT/TTC</td>
<td>Phe/Leu</td>
<td>0.16</td>
<td>NS</td>
</tr>
<tr>
<td>SNP158</td>
<td>ACA/GCA</td>
<td>Thr/Ala</td>
<td>0.42</td>
<td>NS</td>
</tr>
<tr>
<td>SNP174</td>
<td>ACT/ACG</td>
<td>Thr</td>
<td>0.13</td>
<td>S</td>
</tr>
<tr>
<td>SNP198</td>
<td>CTC/CTA</td>
<td>Lec</td>
<td>0.17</td>
<td>S</td>
</tr>
<tr>
<td>SNP224</td>
<td>CTT/CTA</td>
<td>Lec</td>
<td>0.17</td>
<td>S</td>
</tr>
<tr>
<td>SNP224, SNP224</td>
<td>GCT/GCA</td>
<td>Gly/Arg</td>
<td>0.33, 0.33</td>
<td>NS, NS</td>
</tr>
<tr>
<td>SNP226, SNP226</td>
<td>GCC/GCT</td>
<td>Gly/Cys</td>
<td>0.33, 0.33</td>
<td>NS, S</td>
</tr>
</tbody>
</table>

The Dn/Ds (non-synonymous:synonymous) ratio, 1.4, is considerably higher than 1.0, the expected ratio under neutrality, suggesting the cvSI-1 gene has been under positive selection. SNP223 and 224 occurred together at the same codon, both are non-synonymous changes (Table 2). SNP226 and 228 are also found at the same codon: SNP228 is synonymous and SNP226 is non-synonymous.

Six representative SNPs were identified with HAPLOVIEW software (http://www.broad.mit.edu/mpg/haploview/) and subsequently considered for HRM assay design and possible genotyping. Among the 6 SNPs, four had narrow flanking sequences and were not suitable for primer design. HRM primers were designed for two SNPs: 94 (non-synonymous) and 198 (synonymous). HRM assay for SNP94 did not work well, and genotypes could not always be determined with certainty. HRM assay for SNP198 produced clear and reproducible genotypes (Fig. 2). It was subsequently genotyped and characterized in selected families and populations.

3.2. Association with disease resistance

To determine if variation at cvSI-1 is associated with disease resistance, SNP198 was genotyped in three families before and after disease-caused mortalities. Mortalities bracketed by the before and after samples were 79.7% for XG3, 74.9% for HB2 and 79.0% for BCMF. Although we could not determine the exact cause of death for each oyster, routine disease monitoring at the experimental site indicated that most of the mortalities were due to heavy Dermo infection. MSX was present but at low levels. Nevertheless, we cannot completely eliminate MSX and its interaction with Dermo as potential causes for some of the mortalities. Clearly, most of the mortalities were caused by Dermo, but some may be caused by MSX and other unknown factors such as summer stress. It is reasonable to assume that after mortality oysters are more resistant to Dermo and possibly also MSX than the before mortality oysters.

In family XG3, SNP198 was not variable. In the HB2 family, parental genotypes were AC × CC. Before mortalities, HB2 progeny consisted of half AC and half CC genotypes as expected under Mendelian inheritance. After mortalities, the frequency of CC genotype increased from 48.9% to 63.8% (Table 3). The shift in genotype frequency was statistically significant (p = 0.0336). The frequency of allele C increased from 74.5% to 81.9%, and the increase was marginally significant with a p-value of 0.0607. In the BCMF family, parental genotypes were AC × AC. The progeny consisted of 29.2% AA, 53.8% AC and 17% CC before mortalities. After mortalities, the genotype frequency shifted to 14.9% AA, 57.4% AC, and 27.7% CC. The shift in genotype frequency was significant (p = 0.0107). The shift in allele frequency was also significant (p = 0.0126). The frequency of the C allele increased from 43.6% prior to mortality to 56.4% after the mortalities. The post-mortality increase in frequencies of the CC genotype or the C allele suggests that the C allele is associated with improved survival or disease resistance.

To determine if resistance of the C allele had led to its enrichment in disease-resistant strains, we genotyped SNP198 in disease-resistant (NEH) and susceptible (FMF and UMFS) strains as well as the wild base population (CTW) from which these strains were developed. No significant deviation from Hardy–Weinberg equilibrium (HWE) was observed in all populations (Table 4), suggesting that selection at this locus is at equilibrium. In the wild population, the frequency of the C allele was 69.1% (Table 4). In the two susceptible strains, the C allele frequencies, 50% in FMF and 45.8% in UMFS, were significantly (p < 0.001) lower than that in the base population. In the disease-resistant NEH strain, the C allele frequency was 83.3%, which was significantly higher than that in the wild base population and the two susceptible strains. Oysters of the wild population consisted of 53.2% CC, 31.9% AC and 14.9% AA genotypes. In comparison, the NEH strain consisted of 68.8% CC, 29.2% AC and only 2.1% AA genotypes (Table 4).

We also analyzed body size of oysters with different genotypes in families and populations and found no significant association. Thus, there is no evidence that variation at SNP198 affects body size.

4. Discussion

Dermo and MSX have devastated eastern oyster populations in the mid-Atlantic region for decades. They also pose serious threats to aquaculture development. We have learned much about how these diseases progress in infected oysters and how they affect oysters at the individual and population levels [1,2]. We also know that some oysters can resist these diseases as evidenced by the development of disease-resistant stocks [15,16]. However, we know little about the molecular mechanisms of disease resistance in oysters. The identification of disease-resistance genes should improve our understanding of disease resistance and potentially contribute to the development of disease-resistant oysters.

Innate immune response is the primary defense mechanism in oysters and other invertebrates, where hemocytes and cytokines play important roles. Oyster plasma is known to contain protease inhibitors that may participate in host defense [17]. CvSI-1 is a recently identified SPI from the eastern oyster that may be involved in Dermo resistance. It inhibits both proteases from P. marinus and the proliferation of the parasite in vitro [9,10]. To establish a direct link between cvSI-1 and disease resistance, we focused on polymorphism in the coding region of this gene as
studied its association with improved survival and with disease-resistant strains.

Our data show that cvSI-1 is highly polymorphic. It contains 12 SNPs in the coding region in the panel of oysters that we sequenced. This corresponds to a high polymorphism level of about one SNP per 23 bp. Oysters are well known for their high levels of polymorphism [18]. A previous study in the same species of many genes reported a general polymorphism level of one SNP per 24 bp in coding regions [19], which is similar to what’s observed here for cvSI-1, one SNP per 23 bp. In the Pacific oyster, the polymorphism is estimated to be one SNP per 60 bp in coding regions [20], which is considerably lower than what has been reported for the eastern oyster. It is possible that the low estimate for the Pacific oyster is due to the use of introduced hatchery populations which may not be representative of wild populations. It is also possible that the eastern oyster is more polymorphic than the Pacific oyster. The observation of a greater than 1.0 Dn/Ds ratio suggests that cvSI-1 has been under strong positive selection. This finding fits well with the expectation of a disease-resistance gene since oyster populations along the Atlantic and Gulf coasts have been under strong selective pressure by diseases.

By genotyping one of the SNPs of cvSI-1 in oysters before and after disease-caused mortalities, we show that variation at cvSI-1 is indeed associated with survival of oysters under diseases. The C allele at SNP198 increased in frequency after disease-caused mortalities in both families where the locus was variable.

Table 3
Sample size and percent genotypes and alleles observed before and after mortalities at SNP198 of cvSI-1 in two families of the eastern oyster. p-Values are from chi-square test of before and after mortality frequencies.

<table>
<thead>
<tr>
<th>Family: genotype</th>
<th>Time</th>
<th>Genotypes (%)</th>
<th>Alleles (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n  AA  AC  CC</td>
<td>n  A  C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB2: AC × CC</td>
<td>Before</td>
<td>47  0  51.1  48.9</td>
<td>94  25.5  74.5</td>
</tr>
<tr>
<td></td>
<td>After</td>
<td>47  0  36.2  63.8</td>
<td>94  18.1  81.9</td>
</tr>
<tr>
<td>BCMF: AC × AC</td>
<td>Before</td>
<td>47  29.2  53.8  17.0</td>
<td>94  56.4  43.6</td>
</tr>
<tr>
<td></td>
<td>After</td>
<td>47  14.9  57.4  27.7</td>
<td>94  43.6  56.4</td>
</tr>
</tbody>
</table>
monitoring at the experimental site during the period of this study suggests that most of the mortalities were caused by Dermo, and some might possibly be caused by MSX. Thus, it is likely that the C allele is associated with Dermo resistance. This is in agreement with previous observations that cvSI-1 inhibits proteases of *P. marinus* and the proliferation of Dermo parasites [9,10]. The observation that cvSI-1 had higher levels of expressed in oysters selected for Dermo resistance than in susceptible oysters supports its role in Dermo resistance [10], although the resistant and susceptible oysters were from two very different populations (LA and ME), respectively, and many genes may be differentially expressed between these two populations due to their adaptation to different environments in LA and ME. The fact that there was no change in the expression of cvSI-1 upon Dermo challenge in either group [10] seems to caution against cvSI-1’s involvement in Dermo resistance. Furthermore, it is unknown whether cvSI-1 has similar inhibitory effects on the proliferation of *H. nelsoni*, the pathogen for MSX. Since we could not eliminate MSX as a possible cause for the mortalities observed in this study, it may be prudent to simply state that the C allele is associated with disease resistance. Further studies are needed to clarify if cvSI-1 is associated with resistance to Dermo, MSX or both.

The association of a gene with improved survival in full-sib families does not necessarily mean that the gene is responsible for the observed phenotypes, because in families linked genes within a large genomic region can co-segregate and cause the same phenotypic variation [21]. Only if the association persists in wild populations, we may conclude that the gene is responsible for the observed phenotypes, or it is very closely linked to the causative gene. Association studies in wild populations require large numbers of resistant and susceptible phenotypes, which are difficult to obtain. The analysis of disease-resistant and susceptible strains provides a cost-effective alternative. The disease-resistant strain used in this study has been selected for MSX resistance since 1960 and for Dermo resistance since 1990. It has shown strong resistance to MSX and moderate resistance to Dermo [12,15]. If the C allele at SNP198 is associated with disease resistance, selection over many generations should have enriched its frequency. Our results show that the frequency of C allele is significantly higher in the resistant strain than in susceptible strains and the base population from which all strains were derived. This finding provides strong evidence that the C allele is associated with disease resistance at population level, which means that either cvSI-1 is directly responsible for some of disease resistance or it is closely linked to a causative gene nearby.

SNP198 is a synonymous mutation that does not cause changes in amino acid sequences. It is possible that its association with disease resistance is due to its linkage to a non-synonymous mutation nearby that affects the function of cvSI-1 and host defense. Since expression level of cvSI-1 is higher in Dermo-resistant oysters [10], it is plausible that the C allele at SNP198 is linked to a polymorphism at the promoter region that increases the expression of the cvSI-1 gene. It is also possible that synonymous mutations are not silent and may cause functional changes in the protein. Synonymous mutations may involve rare codons that affect the timing of co-translational folding and therefore the structure and function of the protein [22,23]. It is unfortunate that we don’t have the sequence for the promoter region at this time and non-synonymous SNPs in the coding region could not be genotyped in this study. Further studies should be directed to the discovery of SNPs in the promoter region and interrogation of all SNPs associated with this gene. The identification of the functional polymorphism that is responsible for disease resistance will shed light on the molecular mechanism of gene action for cvSI-1.

This study also shows that HRM analysis can be used for simple and effective SNP genotyping in oysters. SNPs can be genotyped with a variety of methods ranging from restriction enzyme digestion, sequence-specific conformational polymorphism, allele-specific amplification and single-base extension [24]. Most of these methods are either too expensive or too complicated for use in non-model species such as oysters. Recently, two real-time PCR based genotyping methods have been developed and shown some promise. One is the Tm-shift assay that is based on amplification of two alleles with primers of different length [25]. The Tm-shift method has been successful used in SNP genotyping in oysters [19,26]. It requires three primers whose ratio has to be optimized. The other is the HRM assay which can detect small changes in melting curve caused by an SNP in a short amplicon [27]. It is a simple PCR with two regular primers targeting a short sequence. This is probably the first time that HRM is used in oysters and molluscs. Our results show that HRM assay can provide accurate genotyping of SNPs in the eastern oyster. The HRM assay for SNP198 produced clear and reliable genotyping results (Fig. 2), which are supported by sequence data. However, optimization of primers and PCR conditions are clearly needed in some cases such as SNP94 which we were not able to genotype unambiguously with HRM. Nevertheless, our study shows that HRM can provide a simple and effective method of SNP genotyping in oysters. The SNP198 HRM assay developed here should be useful for future genetic analysis of populations and marker-assisted selection.

In conclusion, this study provides direct evidence that variation at cvSI-1 is associated with disease resistance in the eastern oyster. The HRM assay developed here provides a useful marker for further studies on this important gene.

**Acknowledgements**

We thank Paul Rawson and Frank M. Flower’s Oyster Company for providing some of oysters used in this study. We thank Gregory DeBrosse, Yongping Wang, Lingling Wang, Liusuo Zhang and Hong Yu for help with oyster culture and sampling. This study was conducted at Rutgers University and supported in part by grants from the NOAA Sea Grant Oyster Disease Research Program, NSF EID program and a China Oversea Scholarship. This is publication IMCS-2011-10 and NJSG-11-788.

**References**


