

Sequence variability of fibrinogen-related proteins (FREPs) in *Crassostrea gigas*

ZHANG LinLin, LI Li^{*} & ZHANG GuoFan

Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

Received October 19, 2011; accepted March 20, 2012; published online May 22, 2012

Fibrinogen-related proteins (FREPs) are lectins with at least a fibrinogen-like domain. They are widespread in Mollusca and play an important role as immune pattern-recognition receptors. These diverse molecules provide interesting models for the study of the evolution of innate immunity. Although oysters are among the most studied mollusks, there have been few investigations of this gene family in oysters. This paper reports the cloning of the first oyster FREP, named CgFREP, from the Pacific oyster *Crassostrea gigas*. Five types of FREP (CgFREP-1 to CgFREP-5) were obtained based on a single pair of primers, confirming their high diversity. The proteins deduced from these FREPs contained 253–363 amino acids and were divided into three subfamilies based on sequence identities (cutoff point <86%). The topological structure of the phylogenetic tree showed the five CgFREPs clustered together and were completely separated from the subgroup of FREPs obtained from *Biomphalaria glabrata* (BgFREPs). This implies that FREP diversity evolved before the BgFREPs branched off. An interesting aspect of the CgFREPs was their individual hypermutation. Ninety-three SNPs and five indels were characterized among CgFREPs. Some mutations (e.g. an insertion in CgFREP-5) could cause premature translation termination. Multiple alignments suggest recombination events might have occurred during exchanges between alleles, or by somatic diversification, or both. Individual sequence variability parallels the extraordinary complexity of the innate immune system, suggesting a potential capacity to recognize and eliminate many kinds of pathogens, even in a primitive system. The diversity of immune pattern-recognition receptors may offer an alternative strategy that functions in the absence of highly specific immune globulins, which only emerged in vertebrates.

***Crassostrea gigas*, FREPs, fibrinogen-related proteins, recombination, oysters**

Citation: Zhang L L, Li L, Zhang G F. Sequence variability of fibrinogen-related proteins (FREPs) in *Crassostrea gigas*. Chin Sci Bull, 2012, 57: 3312–3319, doi: 10.1007/s11434-012-5155-6

Maintenance of the immunological integrity of organisms is a driving force in evolution. A large number of heterogeneous non-self recognition strategies and defense mechanisms contribute to adaptive immunity. It is generally believed that only jawed vertebrates have a well-developed adaptive immune system. In contrast, innate immune mechanisms share similarities among invertebrates, vertebrates, and even plants [1,2]. Besides the rearrangement of immune globulin (IG) and T-cell receptor systems in adaptive immunity, another mechanism for the generation of immune diversity has emerged over the last several decades, i.e. a high genetic polymorphism of innate immune receptors [3]. Unlike adap-

tive immune mechanisms, the recognition specificity of innate immunity is germline encoded. Recent findings suggest that somatic mechanisms of receptor diversification are analogous to those of adaptive immunity, e.g. variable lymphocyte receptors in jawless vertebrates [4,5], extensive alternative splicing of the Down syndrome cell adhesion molecule (DSCAM), and fibrinogen-related protein (FREP) receptors in *Biomphalaria glabrata*.

FREPs are a family of lectins containing a fibrinogen-related (FBG) domain. This gene family consists of various proteins, including tenascins, tachylectins, ficolins, angiopoietins, ixoderins, and fibrinogen β and γ chains. FREPs have been universally identified in vertebrates [6], urochordates [7] and invertebrates [8,9] and have been reported as

*Corresponding author (email: lili@qdio.ac.cn)

pattern recognition proteins in invertebrate animals [8,9]. To date, only *B. glabrata* and *Aplysia californica* have been reported to possess lectins with the specific IG and FBG domains that have been proposed to play an important role in receptor diversification. FREPs can be up-regulated under relatively specific immunostimulation [10] and they are able to bind to and precipitate parasite antigens [9,10]. The significant role of FREPs in innate immunity implied by these research reports is widely recognized.

High levels of sequence diversity are characteristic of FREPs and may contribute to the innate immunity system of invertebrates [11,12]. The biological function of this diversity has been addressed in some cases. Multiple variants of Toll-like receptors from the sea urchin *Strongylocentrotus purpuratus* and FREPs from the mosquito *Anopheles gambiae* were reported to be able to recognize and differentiate bacteria and fungi [13,14]. Therefore, multiple variants accumulating around a unique origin may play important roles in microbial target specificity, which is one of the most important aspects of the evolutionary arms race between pathogens and their hosts.

The pacific oyster, *Crassostrea gigas*, is a widely used model for molluscan studies because of its economic, ecological and evolutionary importance [15]. The species inhabits coastal tidal zones where it is exposed to unstable conditions and diseases caused by parasites, viruses and bacteria. Such environmental stress would impose an evolutionary pressure for the development of sophisticated defense mechanisms.

This study reports a further remarkable example of highly diverse immune molecules in *C. gigas* and represents the first cloning of an oyster FREP containing a complete open reading frame. In addition, we have investigated the diversity of FREPs in this oyster and have demonstrated the existence of a high diversity of mutations in the transcripts. Such hypermutation could potentially change the structure of proteins and perhaps change the function of the genes. We infer that this sequence diversity results from recombination between alleles, or from somatic diversification, or both.

1 Materials and methods

1.1 Animals

Pacific oysters *C. gigas* of mean shell height 110 mm were obtained from an oyster farm in Qingdao, China, and acclimatized for one week in seawater tanks at (18±1)°C and salinity 30 ppt.

1.2 Database searching and sequence retrieval of FREPs in *C. gigas*

Molluscan FREPs cloned in *Argopecten irradians* (ACB29735) and *Biomphalaria glabrata* (Q95UV8, Q8WQX8, Q95UV9,

Q8WQX7, Q86H05, Q86GZ8, Q86H06, Q86H02, Q86GZ9 and Q86H03) were downloaded from the GenBank database. Using TBLASTN, these homologous protein sequences were screened against the *C. gigas* EST database, which was also downloaded from GenBank. An *E*-value of 1×10^{-5} was used and only ESTs longer than 100 bp were considered. The filtered ESTs were then manually assembled and aligned with homologous FREPs of *A. irradians* and *B. glabrata*. Two ESTs (EW778094 and EW778622) were selected because they aligned with the 5' and 3' ends, respectively, of homologous FREPs. Groups of primers were designed. FREPF1 (5'-GACTTGATATCGGTTCACAGCATCX-3') and FREPR1 (5'-CTGCTACATTGGCAACATAATGAGGA-3') amplified the longest PCR product and were used in subsequent studies.

1.3 Cloning of groups of FREPs

Total RNA was isolated from gills using Trizol reagent. PCR amplification of *C. gigas* FREP based on the gill tissue of a single oyster was conducted and PCR products from 1 to 2 kb in length were isolated and cloned into the pMD18-T vector (Takara Bio Inc., Japan). Colony PCR was conducted with the primers FREPF1, FREPR1, FREPF2 (5'-CGGATAACTCGGCAATGGGTAGACT-3') and FREPR2 (5'-CCGTTCTTGTATTCCGTCCAGTTCC-3'). FREPF2 and FREPR2 were used as nested primers to better select different transcripts of FREPs.

1.4 Analysis of nucleotide and deduced amino acid sequences of CgFREPs

Homology searches were performed with a local BLAST package. cDNA and deduced amino acid were analyzed using the software package DNAMAN. Deduced protein domains and N-linked glycosylation sites were predicted in the online analysis of the ExPASy Proteomics Server. Calculated molecular masses and theoretical isoelectric points were predicated by a protein molecular weight and composition calculator (http://www.proteomics.com.cn/proteomics/pi_tool.asp). Multiple protein sequence alignment was created with ClustalX 1.83 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). The phylogenetic tree of selected CgFREPs was constructed with the MEGA program (version 4.0.2) using the neighbor-joining method and the reliability of the branching was tested using bootstrap resampling (with 1000 pseudo-replicates).

2 Results and discussion

2.1 cDNA and the deduced amino acid of CgFREP-1

cDNA of FREP with a complete open reading frame (ORF) was acquired and designated as CgFREP-1. The ORF of 1106 base pairs (bp) encoded 275 amino acid residues

(Figure 1). The length of CgFREP-1 was 1205 bp and contained a 5'UTR (untranslated region) of 53 bp, and a 3'UTR of 45 bp that included a stop codon TGA. CgFREP-1 contained an adenine three bases upstream of the start codon ATG, plus another adenine followed by another start codon ATG, which is consistent with Kozak's rule. Additionally, a stop codon (TAA) was present upstream of this ATG in the same coding frame of CgFREP-1. These two features suggested that ATG (53–55) in CgFREP-1 was the initiation methionine for translation. The SMART database predicted

```

1      GACTTGATATCGGTTACAGCATCGTAAGCGATCTGACGGATAACTCGGCAATGGGTA
1
2      M G
61     GACTGTTAGGTATTCTGATTTGACTTCATTGCTTGATCACTTAATGGACAGGTGTTG
3      R L L G I L I L T S L L V S L N G Q V L
121    TTGATGATATTGGAGATAGACAGAGGTCAACAGACTTGGCCTAACATGCATGAGAACGG
23     F D D I G D R Q R S T D L A L Q C M R T
181    GGATTCAAGCAATTAGTAGCAAGGTATAGAAAAGATGGTAAACGAAAAAATGGAGAGTCG
43     G I Q Q F S Q G I E K M V T K K L E E V
241    AAAACGTCATCGAGACAAAGATAACCAATTGCTAGAAAAGATGGTAAACGAAAAATCGTATC
63     E N V I E T K I T N L S R K V E K I S D
301    TTATTACACAACCTAGACCCGAGCAGGACTCTAAACGAGAAATGGACTTTTACTT
83     L I Y T T Q T E Q R L L K Q E M E S F T
361    CTAAAACTGACACTATGTTAACAAATTGTGAGACAAAGATCGCAATTGTCAGAAATG
103    S K T D T M L N N C E T K I S N L S R N
421    TTGAGAAAATCAGTGTATATTACACAAACTCAGACCCGAGCAGCAGCTCTAAAGCAGG
123    V E K I S D H I T T T Q T E Q R L L K Q
481    AAATGGAGCTTTAAACTCAAAGCTGTTCTGGTAACAAAAGTCATAACAAACTGTG
143    E M E S L N S K A D F L V N K S L T N C
541    CTGACATTTGAACATTATCCACATACCAGAGGGAGGGATGGTGTATAATATCTTG
163    A D I L N I Y P H T R G R D G V Y N I F
601    ATTCAAAGGCCGTGACTGTGATATGACCACTGAGAACGGAGGATGGACGGTAGTCAGC
183    D S K A V Y C D M T T E N G G W T V V Q
661    GGAGAGTCATGGATCAGTGGACTTTACCGTAACGGACAGAACGGTTTG
203    R R V N G S V D F Y R N W T E Y K N G F
721    GATTGCTGATCACAGGTTGGATAGGAACGACATGCTACATAGACTGACATTACTGA
223    G F A D H E F W I G N D M L H R L T L L
781    AGCCCCAGGAACCTCGATTGACATGGAGAGATTAATGGAGAGAACGCTATGCTGTAT
243    K P Q E L R F D M E R F N G E K A Y A V
841    ACTCTAGTTTCTGTCGGGACGAGGCCAGTAATATCAGTCACAGGTGACTGGATACA
263    Y S S F S V G D E A S K Y Q L Q V T G Y
901    GTGAAATGCAGGAATAGCCTGAATAACAACAACATGAAGTCTCTACACCGGACC
283    S G N A G N S L N N N N N M K F S T P D
961    AGGATAACGACGGGAACAGTTAACACTGTGCCACCAACAGGAGATCGGCATGGTGTCA
303    Q D N D G N S Y N C A T N R R S A W W F
1021   ACAGCTGTACCCCTACCAACCCATAATGGAGAGTACACGGACTCTGAGAAAATGGTGG
323    N S C T L T N P N G E Y T D S E K T G W
1081   CATACTATTGTTACTGGAAAAACTCGAGGATATCTCTGAAGTCATAACAGCTGA
343    T Y I N W Y H W K N S R I S L K S I Q L
1141   TGATTGCTCCCGAGCCTGAATCGAGATGTAACGTTACCTCATTATGTTGCCAAATGT
363    M I R P R A *
1201   AGCAG

```

Figure 1 Nucleotide and the deduced amino acid sequences of CgFREP-1 of the pacific oyster *Crassostrea gigas*. The start codon (ATG) and the stop codon (TGA) are boxed. The potential functional FBG domain predicted by ScanProsite is marked in grey. Putative glycosylations sites are underlined.

that CgFREP-1 had a similar domain structure to that of some invertebrate FREPs [9,11,14]. They all contained a C-terminal FBG domain. The calculated molecular mass of the deduced CgFREP-1 was 43.4 kD, and the theoretical isoelectric point (pI) was 6.55. To survive infection in the absence of adaptive immunity, invertebrates rely completely on innate immune defense mechanisms to eliminate invading pathogens. A significant role for FREPs in immunity has been described [9,16]. The oyster is a representative molluscan species but, previously, no FREP containing a complete open reading frame has been reported for oysters. In the present study, five FREPs were cloned from the Pacific oyster. These sequences should provide useful information for future FREP studies.

2.2 Groups of FREPs

With the primers located at the 5'UTR and 3'UTR of CgFREP-1, colony PCR products exhibited high diversity. A pair of nested primers (FREP1F/R1) was designed to help identify different CgFREP transcripts (Figure 2). We distinguished five variants of CgFREP from lanes (1,3,6), (2,4,9), 5, 7 and 8. The variant identical to that cloned was designated as CgFREP-1 and the other four FREP transcripts were named CgFREP-2, CgFREP-3, CgFREP-4 and CgFREP-5 (Table 1). These four FREP cDNAs were all composed of complete open reading frames and the deduced translated peptides varied in lengths from 253 to 363 amino acids.

The nucleotide sequences of *C. gigas* FREP-2/3/4/5 were calculated as 82%, 93%, 73% and 76% identical with CgFREP-1, respectively. Thus, five CgFREPs could be divided into three subgroups with identity cutoff points of <86% [17]. Compared with subgroup 1 (CgFREP-2 and CgFREP-4), a 138 bp insertion existed in subgroup 2 (CgFREP-1 and CgFREP-3). CgFREP-5, belonging to subgroup 3, had the lowest molecular mass because the insertion in CgFREP-5 changed the coding frame, causing a TAG codon to occur at position 760, and early termination of the translation. Multiple alignment of nucleotide residues showed that many regions were highly conserved among the five CgFREPs (Figure 3).

Basic physicochemical and comparative characteristics of the deduced CgFREP proteins are presented in Table 1. All of these protein sequences contained a C-terminal FBG

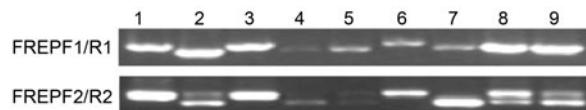


Figure 2 Colony PCR products of *C. gigas* cDNA on a 1% agarose gel. The PCR products on the upper gel were obtained with the primer FREPF1/R1, while the products on the lower gel were obtained with the primer FREPF2/R2. Differences among the bands on the gel help to distinguish the different transcripts of CgFREPs.

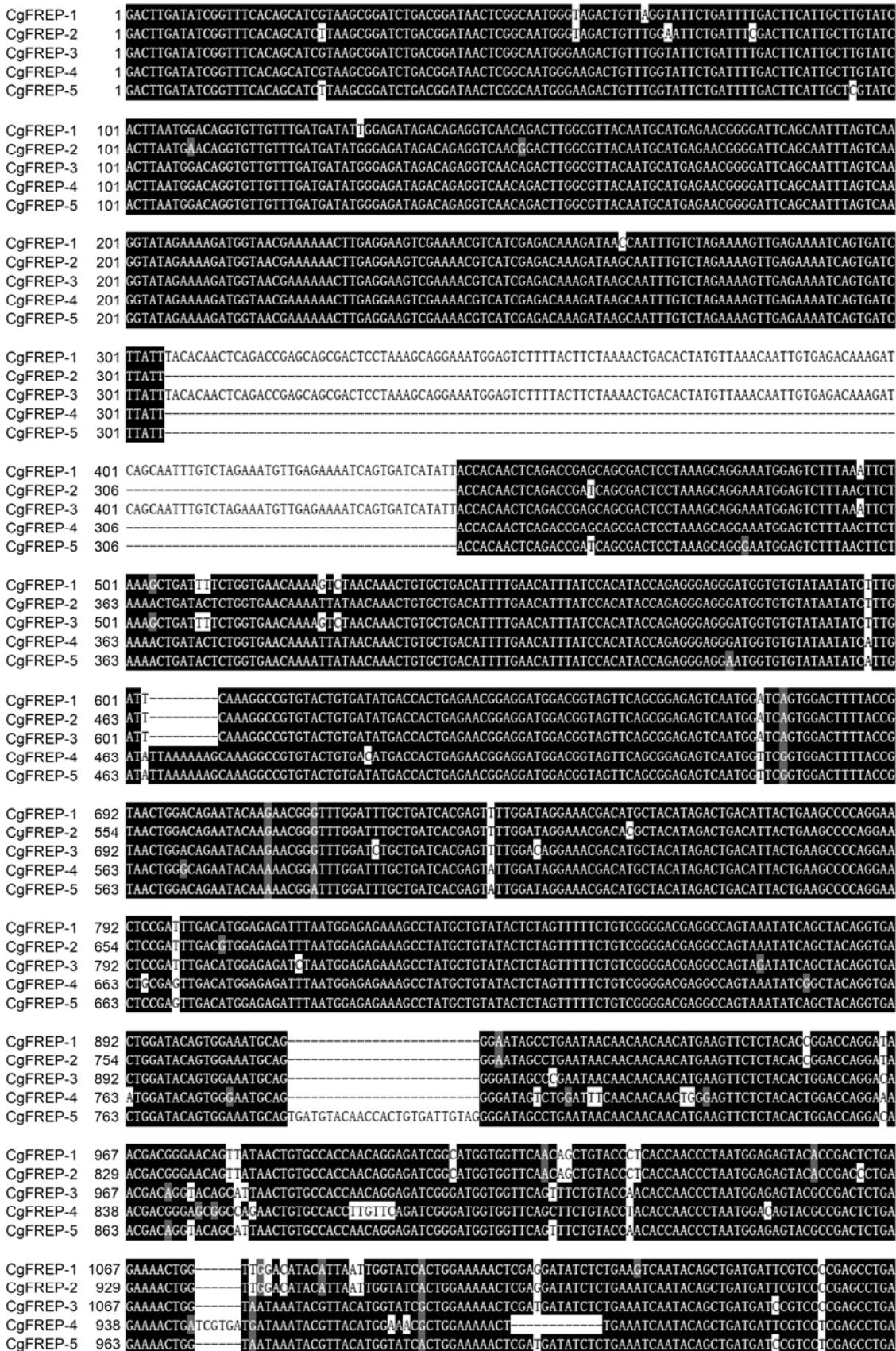


Table 1 Characteristics of *Crassostrea gigas* fibrinogen-related proteins^{a)}

Gene	Length (aa)	MW (kD)	pI	SMART FBG (<i>E</i> -value)	Identities % (CgFREP-1)
CgFREP-1	368	42.4	6.55	1.18×10^{-88}	100
CgFREP-2	322	37.0	6.46	7.26×10^{-89}	82
CgFREP-3	368	42.3	6.91	2.34×10^{-83}	93
CgFREP-4	323	37.2	6.37	2.48×10^{-73}	73
CgFREP-5	253	28.8	7.61	6.72×10^{-20}	76

a) pI, isoelectric point; FBG, fibrinogen-related domain.

domain. The FBG domains were highly conserved. SMART library searches classified these proteins as having FBG domains with *E*-values $<10^{-70}$, except for CgFREP-5 (Table 1). The theoretical isoelectric point of CgFREP-5 predicted protein was also markedly different from the others.

2.3 Homology and phylogenetic analysis of CgFREP proteins

To trace the evolutionary lineage of the five cloned CgFREPs, a phylogenetic tree was constructed based on the amino acid sequences of selected FREPs using the neighbor-joining method (Figure 4). The FBG-containing proteins could essentially be divided into three lineages: ficolins and ficolin-like proteins; tenascins; and fibrinogens [18]. The vertebrate tenascins are closely related to each other and form a separate subgroup, which suggests that vertebrate FBG domain-containing proteins evolved and diversified into a novel lineage. It has been proposed that the primary role of FREP in vertebrates is coagulation, and in invertebrates is defense [16]. The evolutionary diversification of vertebrate and invertebrate FREPs shown in Figure 4 is consistent with their functional diversification. The topological structure of the tree obtained in this study is similar to that of other reports [18]. The CgFREPs and BgFREPs were clustered into separate sister subgroups. The hypermutation of BgFREPs has been well studied. Their sequence contains both FBG and IG domains. It has been demonstrated that this unique domain architecture is an evolutionary novelty that appeared and evolved within only one branch of the protostomes, the heterobranch mollusks (Pulmonata and Opistobranchia) [19]. The individual variability found among *C. gigas* CgFREPs suggests this property of hypermutation existed before the divergence of CgFREPs and BgFREPs.

2.4 Individual variability

An assessment of polymorphisms and individual variability among the five different CgFREPs was carried out (Figure 2; Figure 5(a)). Ninety-three positions were found to be single-nucleotide polymorphisms (SNPs), while five positions were insertions or deletions. These sequence variations were not uniformly distributed. As illustrated in Figure 5(a), SNPs were much more frequent at positions 601–1144 than

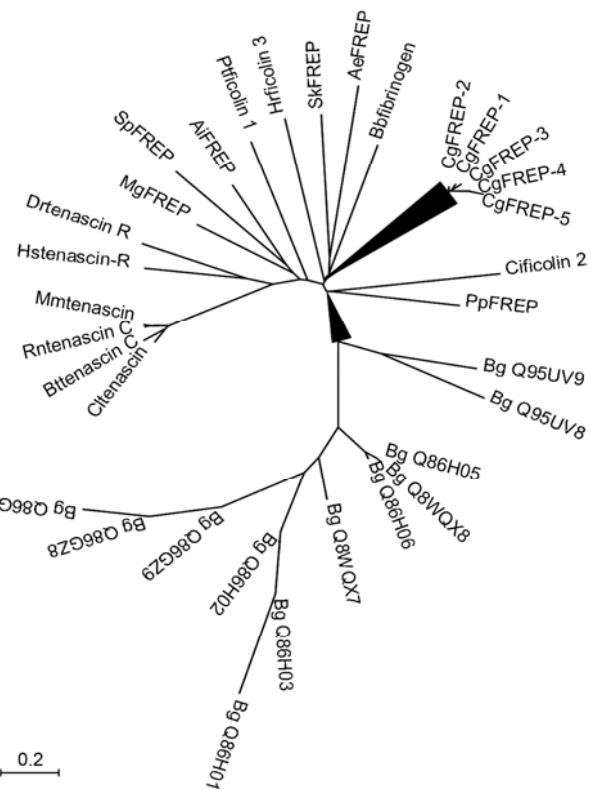


Figure 4 Neighbor-joining tree of proteins containing FBG domain sequences generated by the MEGA program. The scale bar indicates a branch length of 0.2. The sequences shown are as follows: *Canis lupus familiaris* (Cltenascin, NP_001182078), *Rattus norvegicus* (Rntenascin C, NP_446313), *Branchiostoma belcheri* (Bbfibrinogen, ABF8355), *Mus musculus* (Mmtenascin, NP_035737), *Strongylocentrotus purpuratus* (SpFREP, XP_786274), *Ciona intestinalis* (Cifolin 2, XP_002122363), *Bos Taurus* (Bttenascin C, NP_001071494), *Homo sapiens* (Hstenascin-R, CAA91947), *Saccoglossus kowalevskii* (Skfrep, XP_002736868), *Halocynthia roretzi* (Hrtfcolin 3, BAB60706), *Mytilus galloprovincialis* (MgFREP, DQ55822), *Danio rerio* (Drtenascin R, XP_002660896), *Pan troglodytes* (Ptfcolin 1, XP_001170643), *Aedes aegypti* (AeFREP, XP_001662204).

that at positions 1–600. This suggests that the two sections of the FREPs may have different recognition functions, or different protein domain structures, that are subjected to different evolutionary selection pressures.

To clarify whether the nucleotide divergence causes changes in protein structure and function, multiple sequence alignments were conducted based on the deduced amino acid sequences of the CgFREPs (Figure 5(b)). Approximately

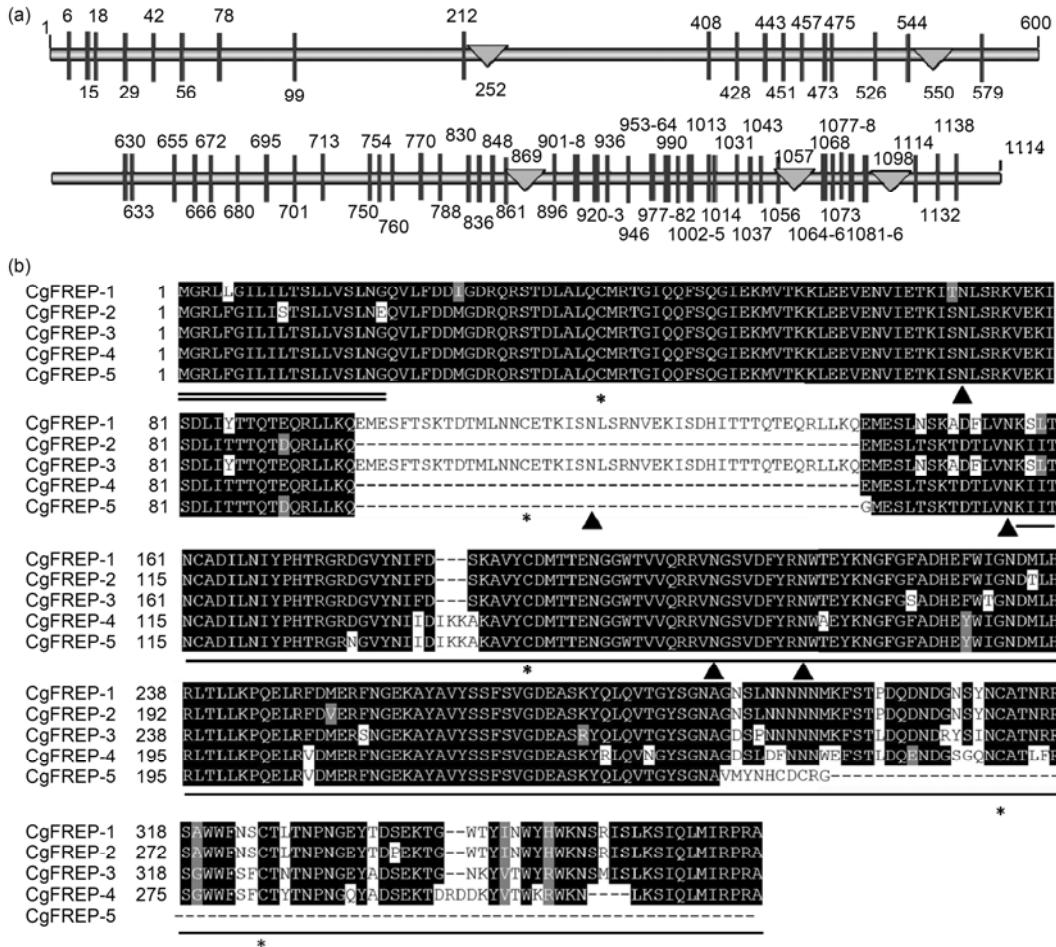


Figure 5 Sequence variability of oyster transcripts of FREPs and their predicted amino acids. (a) Transversal bars represent point mutations in CgFREPs and triangles represent deletions/insertions. Numbers indicate variable positions. The sequences were based on primers FREPF2 and FREPR2. (b) Multiple sequence alignment by Clustal X of the predicted amino acids of the five cloned CgFREPs. Amino acid residues that are conserved in at least 50% sequences are dark-shaded. Asterisks indicate cysteines that are likely involved in the intra-chain disulfide bonds. Putative glycosylation sites are marked with ▲.

one third of the CgFREP-5 FBG domain was lost because the 25-bp insertion changed the reading frame and caused termination mutation. Five cysteines were conserved and were probably involved in the formation of intra-chain disulfide bonds. Five putative glycosylation sites were predicted and probably functioned as pathogen-associated molecular pattern recognition receptors (PRRs). One cysteine and one putative glycosylation site were located at an insertion/deletion of 46 amino acids. These mutations may change the protein structure and perhaps the functions of these genes.

The evolution of immunity is a widely studied topic. Investigations into the acquisition of complexity are providing glimpses into as yet unforeseen mechanisms of immune receptor diversification. Among invertebrates, the Toll pathway and RNA processing of Dscam in the fly, hypermutation of the gene family FREP in the snail, expansion of Toll-like receptors in the sea urchin, and alternative V regions of VCBPs in the amphioxus, are four important current ideas of alternative model systems for generating immune diversity [20]. The hypermutation of FREP in snails

has attracted widespread attention. However, few studies have focused on FREPs of other mollusks. Recently individual variations of FREPs were reported in *Mytilus gallo-provincialis* [18]. The individual variation found in *B. glabrata* and *M. galloprovincialis*, and in *C. gigas* in this paper, all point to the extraordinary complexity of invertebrate immune systems. This may offer an alternative strategy that assists survival in the presence of a range of pathogens, even in the absence of highly specific IgGs [21].

To further investigate the mechanism of high and variable sequence diversity of *C. gigas* FREPs, the possibility of recombination events was assessed by visual inspection of the alignments. This analysis suggested past recombination events between a number of different groups. As shown in Figure 6, CgFREP-1 and CgFREP-3 were identical while CgFREP-2/4/5 formed a separate group with respect to the first three sections (labeled I, II, and III). In section IV, CgFREP-2 separated from the group CgFREP-2/4/5 and joined the CgFREP-1/3 group. Then, in section V, CgFREP-1/2/3/4 were similar, and an insertion occurred in CgFREP-5.

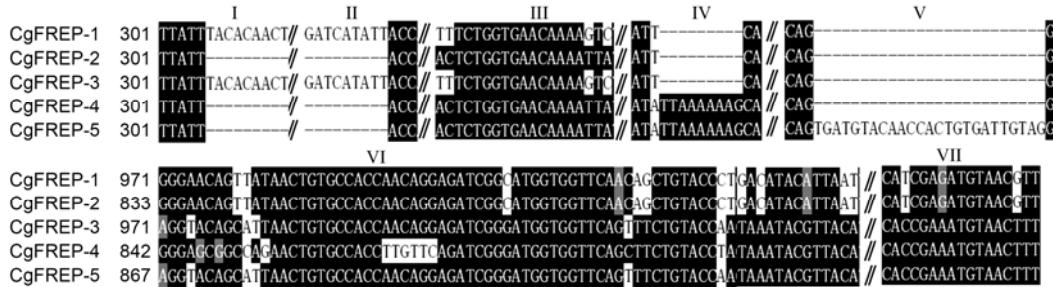


Figure 6 Recombination among CgFREPs. The black shaded regions indicate positions where all sequences share the same amino acid residue. Gaps are indicated by dashes to improve the alignment. Seven sections of the CgFREP sequences were selected. These are separated by the symbol “//” and numbered I to VII.

In sections VI and VII, the CgFREPs were divided into three subgroups, (CgFREP-1, CgFREP-2), (CgFREP-3, CgFREP-5) and CgFREP-4. All of these variations imply past recombination events among the five cloned CgFREPs. Such recombination in immune-related genes also has been inferred in an antimicrobial study of *C. gigas* [22].

In snail studies, 45 and 37 different variants of BgFREP3 were obtained and their strong hypermutation deduced from the absence of insertion or deletions, synonymous nucleotide substitutions, or lethal mutations. Most importantly, the BgFREPs were reported to be somatically diversified [8,23]. Somatic diversification of IGs is well known but has only been reported in vertebrates. Therefore, the study on FREPs offers an alternative approach to the elucidation of invertebrate immune systems. In addition, studies in flies have indicated immune genes that directly interact with pathogens (i.e. recognition and effector genes) are clustered in regions of higher recombination rates, and may also undergo hypermutation [24]. Our study has demonstrated individual hypermutation of the FREP gene family in *C. gigas*. Future research should focus on the mechanism, e.g. on whether CgFREPs are clustered together and highly recombinant or whether somatic recombination occurs in the Pacific oyster.

In our study, the first oyster FREP containing a complete open reading frame was cloned from the Pacific oyster. Five FREP transcripts were identified in a single oyster indicating individual hypermutation. Our findings suggest that this sequence diversity might result from recombination between alleles or somatic diversification, or both. Further studies should include investigations of the mechanisms of hypermutation and the role of CgFREPs in the resistance to pathogen infection.

We thank the reviewers for their efforts on this paper. This work was supported by the National Basic Research Program of China (2010CB126401), National Natural Science Foundation of China (40730845), National High Technology Research and Development Program of China (2012AA10A410), the earmarked fund for Modern Agro-industry Technology Research System, and the Taishan Scholar Program of Shandong and Taishan Scholars Climb Program of Shandong.

1 Nurnberger T, Brunner F. Innate immunity in plants and animals:

Emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. *Curr Opin Plant Biol*, 2002, 5: 318–324

- 2 Hemmrich G, Miller D J, Bosch T C. The evolution of immunity: A low-life perspective. *Trends Immunol*, 2007, 28: 449–454
- 3 Litman G W, Dishaw L J, Cannon J P, et al. Alternative mechanisms of immune receptor diversity. *Curr Opin Immunol*, 2007, 19: 526–534
- 4 Pancer Z, Amemiya C T, Ehrhardt G R, et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature*, 2004, 430: 174–180
- 5 Nagawa F, Kishishita N, Shimizu K, et al. Antigen-receptor genes of the agnathan lamprey are assembled by a process involving copy choice. *Nat Immunol*, 2007, 8: 206–213
- 6 Matsushita M, Endo Y, Taira S, et al. A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin. *J Biol Chem*, 1996, 271: 2448–2454
- 7 Kenjo A, Takahashi M, Matsushita M, et al. Cloning and characterization of novel ficolins from the solitary ascidian, *Halocynthia roretzi*. *J Biol Chem*, 2001, 276: 19959–19965
- 8 Zhang S M, Adema C M, Kepler T B, et al. Diversification of Ig superfamily genes in an invertebrate. *Science*, 2004, 305: 251–254
- 9 Zhang H, Wang L, Song L, et al. A fibrinogen-related protein from bay scallop *Argopecten irradians* involved in innate immunity as pattern recognition receptor. *Fish Shellfish Immunol*, 2009, 26: 56–64
- 10 Adema C M, Hertel L A, Loker E S. Evidence from two planorbid snails of a complex and dedicated response to digenetic (echinostome) infection. *Parasitology*, 1999, 119: 395–404
- 11 Gorbushin A M, Iakovleva N V. A new gene family of single fibrinogen domain lectins in *Mytilus*. *Fish Shellfish Immunol*, 2011, 30: 434–438
- 12 Middha S, Wang X. Evolution and potential function of fibrinogen-like domains across twelve *Drosophila* species. *BMC Genomics*, 2008, 9: 260
- 13 Rast J P, Smith L C, Loza-Coll M, et al. Genomic insights into the immune system of the sea urchin. *Science*, 2006, 314: 952–956
- 14 Dong Y, Dimopoulos G. Anopheles fibrinogen-related proteins provide expanded pattern recognition capacity against bacteria and malaria parasites. *J Biol Chem*, 2009, 284: 9835–9844
- 15 Hedgecock D, Gaffney P M, Gouletquer P, et al. The case for sequencing the Pacific oyster genome. *J Shellfish Res*, 2005, 24: 429–441
- 16 Patrick C, Hanington S M Z. The primary role of fibrinogen-related proteins in invertebrates is defense. *Not Coagul J Innate Immunity*, 2011, 3: 17–27
- 17 Zhang S M, Nian H, Zeng Y, et al. Fibrinogen-bearing protein genes in the snail *Biomphalaria glabrata*: Characterization of two novel genes and expression studies during ontogenesis and trematode infection. *Dev Comp Immunol*, 2008, 32: 1119–1130
- 18 Romero A, Dios S, Poisa-Beiro L, et al. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient

- and complex immune recognition models in invertebrates. *Dev Comp Immunol*, 2010, 35: 334–344
- 19 Gorbushin A M, Panchin Y V, Iakovleva N V. In search of the origin of FREP: Characterization of *Aplysia californica* fibrinogen-related proteins. *Dev Comp Immunol*, 2010, 34: 465–473
- 20 Litman G W, Cooper M D. Why study the evolution of immunity? *Nat Immunol*, 2007, 8: 547–548
- 21 Zhu Y, Ng P M, Wang L, et al. Diversity in lectins enables immune recognition and differentiation of wide spectrum of pathogens. *Int Immunol*, 2006, 18: 1671–1680
- 22 Schmitt P, Gueguen Y, Desmarais E, et al. Molecular diversity of antimicrobial effectors in the oyster *Crassostrea gigas*. *BMC Evol Biol*, 2010, 10: 23
- 23 Hanington P C, Forys M A, Dragoo J W, et al. Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. *Proc Natl Acad Sci USA*, 2010, 107: 21087
- 24 Wegner K M. Clustering of *Drosophila melanogaster* immune genes in interplay with recombination rate. *PLoS One*, 2008, 3: e2835

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.