IOP Conf. Series: Earth and Environmental Science **77** (2017) 012005 doi:10.1088/1755-1315/77/1/012005

Unusual varieties and duplication of Rig-I like receptors encoded in the marine mollusk, *Crassostrea gigas*

Z H Tian and C Z Jiao[1]

Life Science College, Shaoguan University, Shaoguan, 512005, China

E-mail: jiaocz@gmail.com

Abstract. RIG-I like receptors (RLRs) play key roles in sensing non-self nucleic acids in cytoplasm and trigger antiviral innate immune response in vertebrates and human body. Here we carried out *in silico* analysis to identify and investigate the putative RLRs encoded in the genome of marine mollusk, *Crassostrea gigas* (cgRLRs), an invertebrate species. We found the unusual duplication and varieties on domain architecture of putative cgRLRs encoded in the genome of *C. gigas*. Three putative cgRLRs (accessions numbers are EKC24603, EKC31344.1 and EKC38304.1 on GenBank), have the similar domain architecture with that of human RIG-I or MDA5, and one protein (EKC34573.1) with that of human LGP2; The fifth putative cgRLRs (EKC38303.1) is somewhat similar with human RIG-I/MDA5 except that it has only one caspase activation and recruitment domain (CARD) in its N-terminal. Other nine proteins were identified to be partialy similar with RLRs while with the incomplete sequences, which maybe reflect the events of partial duplication of cgRLRs genes occurred in the oyster genome.

1. Introduction

Animals have developed sophisticated mechanisms to detect and combat intrusions of pathogenic microbes in their long evolutionary history. Innate immune receptors in certain cells are responsible for recognizing pathogen-associated molecular patterns (PAMPs), which are structures or chemical patterns that are extensively conserved in bacterial cell wall components, microbial nucleic acids, and certain highly conserved proteins. These innate immune receptors are classically termed patter-recognition receptors (PRRs). Upon their recognition of PAMPs, PRRs trigger downstream signaling that rapidly activate and amplify potent inflammatory antimicrobial responses, mainly through the upregulation of the gene expression and activation of antimicrobial molecules and many cytokines, including NF-κB, type I interferons, or caspases 1 [1].

There are at least three classes of PRRs being identified as sensors of viral nucleic acids: Toll-like receptors (TLRs) which sense viral RNA or CpG-containing DNA in endolysosomes; Viral DNA sensors of structurally diverse group, such as cGAS, IFI16, and DAI; RIG-I-like receptors (RLRs), which sense viral RNA in the cytoplasm of most cell types [2]. In human body, three RLRs members including RIG-I, MDA5, and LGP2 have been identified and studied up to now [2]. RLRs have conserved domain structure. The RIG-I and MDA5 consist of two tandem N-terminal caspase recruitment domains (CARD1 and CARD2) with death domain folds, a DExD/H-box helicase (consisting of two RecA-like helicase domains, Hel1 and Hel2, and an insert domain, Hel2i), and a C-terminal domain (CTD); The third RLR, referred to LGP2, has similar DExD/H-box helicase and CTD conserved domains than that of RIG-I and MDA5, lacks CARD domains in its N-terminal [3,4].

Both RIG-I and MDA5 bind dsRNA through interactions were mediated by the CTD and the DExD/H-box helicase domain, while the different target recognition modes at the CTD allow the two proteins to sense non-self RNAs of different structural patterns [5,6]. The RIG-I senses RNAs bearing a triphosphate (ppp) moiety in conjunction with a blunt-ended, base-paired region at the 5'-end, and it also senses RNAs bearing 5'-diphosphates (5'pp) as shown in a recent study [7]. The CTD of RIG-I caps the duplex terminus through stacking interactions with terminal base pairs [8]. The RIG-I will be active in ATP hydrolysis and downstream signaling after binding of its target RNA [5]. Different from RIG-I, the target of MDA5 is long and accessible RNA duplexes [9]. The CTD of one MDA5 forms a cooperative binding interface with the insertion domain of a neighbor and leads to the oligomerization of MDA5, resulting in formation of long filaments along RNA duplexes [8,10]. The filaments of MDA5 and dsRNA may cause clustering of the CARDs of MDA5, making them active in taking part in immune signaling. LGP2 is present at low levels in the uninfected cells but accumulates in response to virus infection or antiviral mediators, including poly (I-C) and IFNs. Due to absence of CARD or any other signaling interaction domain, the function of LGP2 is largely unknown, while the accumulating evidence shows that it is intimately associated with MDA5 as accomplice and antagonists of antiviral signal transduction [3].

Our understandings of the RLRs are mostly obtained from research on mammals and human. For invertebrates, relatively little is known about their RLRs, if they exist, and their function. The Pacific oyster *Crassostrea gigas* (Thunberg, 1793) is a marine bivalve of the phylum Mollusca, which comprises the largest number of described marine animal species. Many molluscs are important fishery and aquaculture species, and are also models of research in areas, such as neurobiology, biomineralization, ocean acidification and adaptation to coastal environments under climate change [11,12]. Using the genome sequence of *C. gigas* released in 2012 [13], we carried out *in silico* analysis to identify and investigate the putative RLRs coded for in the oyster genome using bioinformatics methods. We found multiple members of putative RLRs, with a variety of domain architectures, encoded within the *C. gigas* genome, which is much more abundant than that in genomes of vertebrates and humans.

2. Methods

2.1. Database mining

To identify putative RLRs (RIG-I, MDA5 or LGP2) encoded in the genome of *C. gigas* (cgRLRs), we employed BLASTP and tBLASTn [14] searches of the human proteins RIG-I, MDA5 and LGP2 from non-redundant protein sequences (nr) and whole-genome shotgun contigs (wgs) database on GenBank. The hits were manually checked based on scores and annotation in GenBank and selected as candidate cgRLRs or genes encoding them to be analyzed further.

2.2. Sequence analysis

The conserved domains of candidate cgRLRs were predicted by searching against the database CDD [15] using a specialized blast (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The full-length amino acid sequences and the conserved functional domains were used in phylogenetic tree analysis. Multiple protein sequence alignments were performed using the MUSCLE [16] programme in MEGA (6.0) [17], which were manually curated if necessary. Phylogenetic trees were constructed using the neighbour-joining method within the MEGA (6.0) package. Data were analyzed using poisson correction, and gaps were removed by pairwise deletion. The topological stability of the neighbour-joining trees was evaluated by 1000 bootstrap replications.

2.3. Genomicstructure analysis

All the genomic data used in this analysis were based on Entrez genome and Entrez nucleotide databases. Protein sequences were taken from Entrez protein databases on National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/).

3. Result

3.1. Sequence identification

With the known human RIG-I protein sequence (NP_055129.2) as query, we performed extensive BLAST analysis of the genome, proteins, and EST databases of the pacific oyster in GenBank. Thirteen proteins annotated as RIG-I family members were identified (table 1).

Table 1. Proteins analyzed in this paper. Blast result is of the human RIG-I protein sequence (NP_055129.2) as query to search the *C. gigas* protein database in GenBank.

GenBank Acc.	Annotation	Max score	Total score	Query cover	E-value	Identity	Length (aa)	DBSOURCE
EKC24603.1	DDX58	360	360	72%	1.00E-106	35%	1198	JH817706.1
EKC38303.1	DDX58	348	348	74%	3.00E-104	34%	948	JH819181.1
EKC31344.1	DDX58	286	286	96%	1.00E-81	28%	1018	JH816862.1
EKC34573.1	DDX58	264	264	75%	2.00E-74	30%	902	JH822697.1
EKC38304.1	IFIH1	238	411	73%	6.00E-65	35%	1260	JH819181.1
EKC37415.1	DHX58	169	169	59%	2.00E-44	27%	538	JH816918.1
EKC34571.1	DDX58	164	164	47%	4.00E-43	30%	440	JH822697.1
EKC29027.1	DDX58	137	137	38%	4.00E-34	30%	459	JH818839.1
EKC24604.1	IFIH1	134	258	46%	3.00E-32	35%	307	JH817706.1
EKC21486.1	Fanconi anemia group M protein	95.1	95.1	56%	1.00E-19	24%	2327	JH817053.1
EKC23131.1	DDX58	69.3	69.3	7%	4.00E-13	46%	158	JH818146.1
EKC29026.1	IFIH1	70.5	70.5	17%	6.00E-13	33%	313	JH818839.1
EKC26346.1	Endoribonuclease Dicer	58.2	103	31%	2.00E-08	31%	1759	JH817198.1
EKC33722.1	Dicer-like protein 2 *	56.2	99	18%	6.00E-08	45%	520	JH819141.1
EKC34572.1	DHX58	49.7	49.7	19%	3.00E-06	25%	270	JH822697.1
EKC31345.1	IFIH1	38.1	38.1	7%	0.018	30%	365	JH816862.1

DDX58, Putative ATP-dependent RNA helicase DDX58, also known as RIG-I
DHX58, Putative ATP-dependent RNA helicase DHX58, also known as LGP2
IFIH1, Interferon-induced helicase C domain-containing protein 1, also known as MDA5
* Also annotated as a putative cgRLRs in this study.

Three related oyster proteins annotated as Fanconi anemia group M protein, Endoribonuclease Dicer, and Dicer-like protein 2 were also included in this study as control due to their high score of BLAST results. We noted that the score of target EKC31345.1 is very low among the listed targets while it is annotated as MDA5. With human MDA5 (NP_071451.2) and LGP2 (NP_077024.2) as query to search, we found the similar targets as above and no other protein was identified annotated as member of RIG-I family. The genomic sequences corresponding with each RIG-I family proteins were also analyzed and will be discussed later in this paper.

3.2. Domain analysis

The conserved domains of putative cgRLRs and human RLRs were predicted by searching in the database of CDD v3.11 with an expect value threshold of 0.01 (figure 1).

Figure 1. Domain architecture of putative cgRLRs and other closely related proteins. (A) human RLRs; (B) putative cgRLRs; (C) Endoribonuclease Dicer (EKC25346.1), Dicer-like protein 2 (EKC33722.1) and Fanconi anemia group M protein (EKC21486.1) of *C. gigas*.

Three putative oyster cgRLRs, which include EKC24603.1, EKC31344.1 and EKC38304.1 have the same architecture as human RIG-I or MDA5, which share all the core protein domains or conserved regions including two Caspase activation and recruitment domains (CARD or DD superfamily), the DEAD-like helicases superfamily (DEXDc or Hel1), helicase insert domain (Hel2i), Helicase superfamily c-terminal domain (HELICc or Hel2), and C-terminal domain of RIG-I (RIG-I_C-RD or CTD). Two other proteins, EKC38303.1 and EKC34573.1, have the same architecture as the three proteins above except that EKC38303.1 has only one CARD predicted, and no CARD predicted on EKC34573.1. The protein EKC34573.1 shares the same domains with human LGP2, while it contains a sequence of about 230 amino acids on the N terminal upstream of the DEXDc/Hel1 domain, which was not predicted as any conserved domains in this study. Three other proteins, EKC37415.1, EKC29026.1 and EKC34572.1 seem to be only the C terminal of RLRs. The proteins EKC37415.1 and EKC29026.1 contain the HELICc/Hel2 and RIG-I_C-RD/CTD domains. The protein EKC34572.1 contains the RIG-I_C-RD/CTD domain only. The proteins of EKC34571.1, EKC29027 and EKC23131.1 were all predicted to contain DEXDc/Hel1 conserved region only. The

protein of EKC31345.1 was predicted to contain two CARD domains. It was surprising that the protein EKC24604.1, which was annotated as Interferon-induced helicase C domain-containing protein 1(MDA5), contains a DEXDc/Hel1 domain, a HELICc/Hel2 domain, and a C-type lectin (CTL)/C-type lectin-like (CTLD) domain (CLECT). The CLECT domain is not contained in the other known RIG-I family proteins. The oyster Endoribonuclease Dicer (EKC26346.1) and Fanconi anemia group M protein (EKC21486.1) have a different architecture each with the RLRs, while they all share the similar domains of DEXDc and HELICc. The protein EKC33722.1 which was annotated as Dicer-like protein 2 on GenBank, contains DEXDc and HELICc domains. We performed the blastp searches with EKC33722.1 as query in human reference proteins database and found that the best hit was human MDA5 (NP_071451.2, with the E- value of 1e-18), so it was also considered as a putative cgRLRs. This was also supported by the following phylogenic analysis.

3.3. Phylogenic analysis

To investigate the relationships of the oyster RIG-I family proteins with those of other vertebrate organisms, the phylogenic tree was constructed with selected available RLRs of vertebrate organisms and the putative cgRLRs (figure 2).

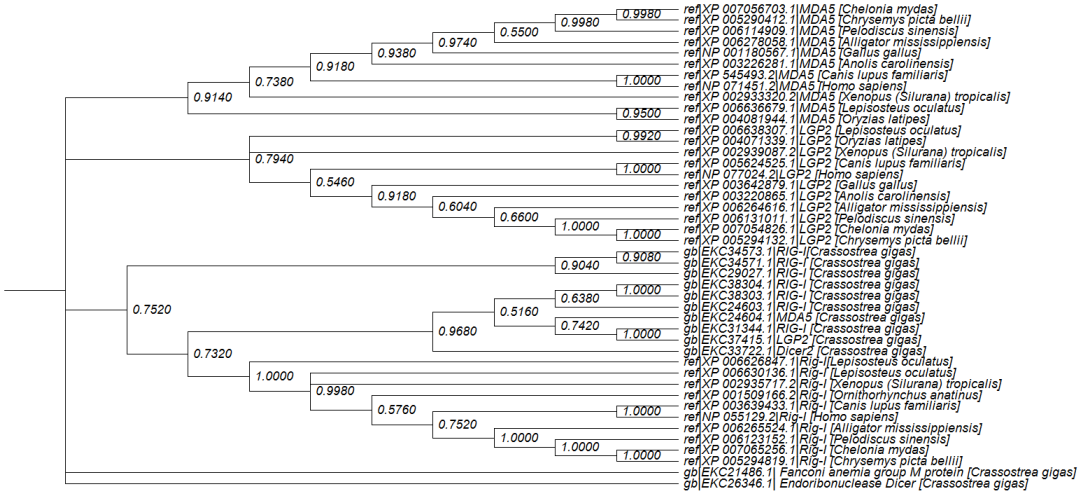

Figure 2. Phylogenetic tree analysis of RLRs.

Three related proteins of *C. gigas,* Dicer-like protein 2, Fanconi anemia group M protein, and Endoribonuclease Dicer were also included in this analysis. All sequences were downloaded from GenBank and aligned with MUSCLE. The trees were constructed using the neighbour-joining method within the Mega 6.06 programme. The bootstrap values of the branches were obtained by testing the tree 1000 times and values over 50% are indicated on the tree.

The phylogenic tree shows that RLRs and cgRLRs were clustered into three groups. Ten of the oyster proteins were clustered into the same group with the vertebrate RIG-I proteins regardless their annotation on GenBank as RIG-I, MDA5, LGP2 or Dicer2. Three other oyster proteins, EKC23131.1, EKC29026.1, EKC34572.1 and EKC31345.1 (figure 1) are not included in this tree due to their sequences being too short. To investigate the relationships of CARD domains in the putative cgRLRs, phylogenetic trees of CARD domain sequences were constructed (figure 3). It is apparent that the CARD of EKC38303.1 and the CARD2 of EKC38304.1, the CARD1/2 of EKC31344.1 and EKC31345.1 are the closest related among the CARDs analyzed (figure 3).

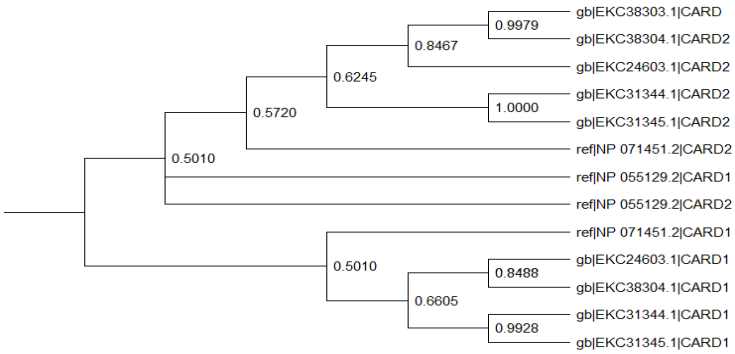

Figure 3. Phylogenetic tree analysis of CARD domains of RLRs.
The sequence of CARD domains of RLRs were predicted by searching in the database of CDD v3.11 with an expect value threshold of 0.01. Sequences of CARDs were aligned with MUSCLE and manually curated. Tree was constructed using the neighbour-joining method within the Mega6.06 programme. The bootstrap values of the branches were obtained by testing the tree 1000 times and values over 50% are indicated on the tree.

3.4. Gene Structure Comparison

Here we investigated the genomic composition of genes encoding cgRLRs, by focusing on the number and location of introns related to the CDS regions that encode different domains (figure 4).

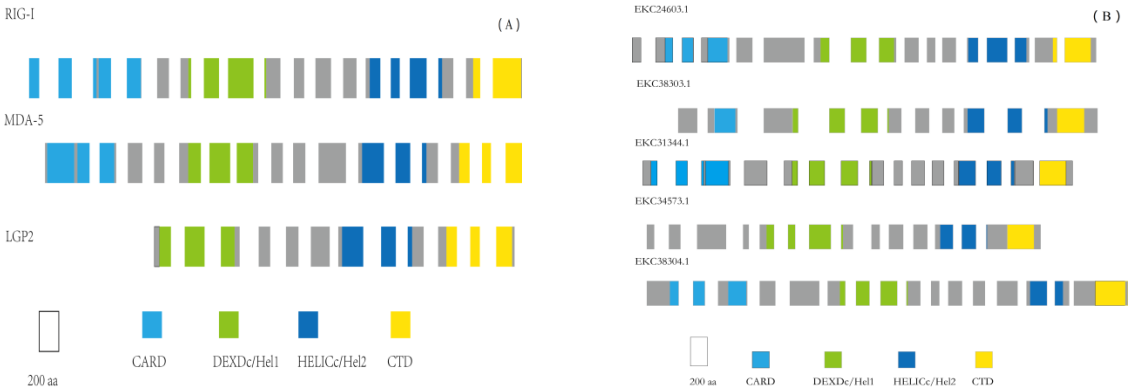

Figure 4. Genomic structure of RLRs in human (A) and *C. gigas* (B).
Major domains are marked with rectangles of different color. The transcript coding region is drawn to scale and introns (blank) positions are illustrated but their size is not drawn to scale.

The gene sequence encoding the CARD1 domain of human RIG-I is split by two introns, while that of human MDA5 is located in one exon. The CARD1 domains of putative cgRLRs are split by one (EKC24603.1, EKC38304.1) or two introns (EKC31344.1). The CARD2 domains of human RIG-I and MDA5 are both split by one intron, while those CARD2 domains of putative cgRLRs (EKC24603.1, EKC38304.1 and EKC31344.1) are all located in one exon. The protein EKC38303.1 contains only one CARD, which like CARD2 domains of other cgRLRs, located in one exon.

The DEXDc/Hel1 domain of human RIG-I is split by three introns, and those of human MDA5 and LGP2 are both split by two introns. For oyster, only the DEXDc/Hel1 domain of EKC24603.1 is split by two introns and this domain of other proteins (EKC38303.1, EKC31344.1, EKC34573.1 and EKC38304.1) are split by three introns. The HELICc/Hel2 domain of human RIG-I are split by 3 introns while that of MDA5 or LGP2 is split by 2 introns. The HELICc/Hel2 domain of cgRLRs is split by 2 introns (EKC24603.1, EKC38303.1, EKC31344.1, and EKC34573.1) or one intron (EKC38304.1). The RIG-I_C-RD/CTD of human RIG-I is split by one intron while that of MDA5 or LGP2 is split by two introns. The RIG-I_C-RD/CTD of oyster protein EKC24603.1 is split by one

intron, other putative cgRLRs (EKC38303.1, EKC31344.1, EKC34573.1 and EKC38304.1) have no intron in the genome sequence encoding their RIG-I_C-RD/CTD domains.

Even though the genome sequence of Oyster is not assembled into contigs corresponding to each chromosome yet, we could still see genes which are vicinity in the genome using the information of their sequence source (table 1). The cgRLRs whose genes are clustered into together in genome are EKC24603.1 & EKC24604.1, which are both located in scaffold1785 (locus JH817706); EKC38303.1 & EKC38304.1, which are both located in scaffold348 (locus JH819181); EKC34572.1 & EKC34573.1, which are both located in scaffold609 (Locus JH822697); EKC31344.1 & EKC31345.1, which are both located in scaffold983 (Locus JH816862); EKC29026.1 & EKC29027.1, which are both located in scaffold1228 (Locus JH818839); EKC34571.1 & EKC34572.1 & EKC34573.1, which are all located in scaffold609 (Locus JH822697).

4. Discussion

The mechanisms of non-self-nucleic acids being recognized by PPRs and triggering signals of IFN production and antiviral responses are evolutionarily conserved from teleost fish to mammals.[18,19]. However, it is generally accepted that IFN-mediated antiviral immunity is absent in invertebrates, as the genes coding IFN homologous and it's major effectors, interferon-stimulated genes (ISGs), are absent in several fully sequenced invertebrate genomes [20, 21]. This view on invertebrate antiviral immunity maybe updated as more studies on antiviral mechanisms in Pacific oyster or other bivalve are available. Earlier studies showed that a STAT-like pathway in *Mytilus galloproviancialis* haemocytes is activated following the treatment with interferon-omega [22]. Many genes encoding homologous to mammalian ISGs, including PKR, IRF, Mpeg1, and IFI44, have been reported to be involved in oyster antiviral response. An "IFN-like pathway" may exist in mollusks [21].

Our research shows that there are at least 14 putative RLRs encoded in the *C. gigas* genome. Seven of them (EKC37415.1, EKC34571.1, EKC29027.1, EKC2313.1, EKC29026.1, EKC34572.1, EKC31345.1, and EKC33722.1) seem to be partial segments of the complete RLRs and contain part of the conserved domains of mammalian (human) RIG-I, MDA5 or LGP2; Three of the putative cgRLRs (EKC24603.1, EKC31344.1, EKC38304.1) have a similar domain architecture to that of mammalian RIG-I or MDA5; The protein EKC34573.1 has no CARD domain and its domain composition and arrangement is similar to that of mammalian (human) LGP2. The protein EKC38303.1 is unique as it contains only one CARD in its N-terminal, while mammalian RIG-I and MDA5 both have two N-terminal CARDs. This type of domains composition is not reported in any other organisms' RLRs yet. Another protein, EKC24604.1, is also interesting in its domain composition, which comprises of a CLECT instead of CTD in its C-terminal. We did not find any other protein with the similar domain architecture in the database of GenBank.

There is no basis to speculate that each of the putative oyster RLRs discussed here has a function in antiviral immunity at present, while the research on their function did emerge recently. The cDNA sequence of a RIG-I homolog in *C. gigas* (cgRIG-I) was cloned and studied [23]. The cgRIG-I sequence deduced from the cloned cDNA (KC702507) has a 100% identity with EKC31344.1 discussed in this study, it is apparent that cgRIG-I and the protein EKC31344.1 are actually derived from the same gene. Zhang's research showed that the mRNA level of oyster RIG-I in haemocytes is increased in response to poly (I: C) treatment. This result implied that the cgRLRs were involved in oyster's antiviral immunity, which is yet to be experimentally verified.

It is unusual that there are so many members of putative RLRs with a variety of domain architectures encoded in the genome of *C. gigas*. Comparatively, there is only one of RIG-I, MDA5 and LGP2 in the genome of human. We have found that in the genome of sea urchin, *Strongylocentrotus purpuratus*, a species of Echinodermata, the unusual high number of putative RLRs does also exist. We speculate that the unusual abundance of putative RLRs in certain organisms represents the state of halfway of the RLRs evolution or it is the result of the animals' adaptation to the specific environment.

Even though the genome sequence of *C. gigas* has not yet been assembled into individual chrosomes, it is can still be inferred that some of the putative RLRs genes are closed to each other in the genome according to their location at the sequence scaffold. It seems that some of the genes are tandem duplicated-at least partialy, and gives rise to the arrangement of the putative cgRLRs genes on scaffolds present in this study. In accord with this assumption, phylogenic analysis shows that those putative cgRLRs whose genes are vicinity in genome are closely related, such as EKC24603.1 and EKC24604.1, EKC38303.1 and EKC38304.1. The protein EKC34572.1 is not included in the phylogenetic analysis, while it is at the same scaffold and proximity to genes of EKC34571.1 and EKC34573.1, which are closed related in the phylogenic tree (figure 2). The phylogenic analysis of CARDs of the putative RLRs gave similar result (figure 3).

The number and location of introns in the putative cgRLRs genes are also investigated here, and in general, are not strictly conserved among the proteins studied. The varieties of cgRLRs are reflected not only by their different domain architecture, but also by the difference in structure of genes encoding them.

The different domains of RLRs play different function. RIG-I exists with a closed inactivation conformation in the absence of dsRNA binding [8]. The helicase and CTD domains recognize and bind viral RNA and release the CARDs, which then recruit and activate the signaling adaptor MAVS (IPS-1) [24]. MDA5 can cooperatively assemble into ATP-sensitive filaments on dsRNA and does not sequester its CARDs [9,10,25]. However, the CTD of MDA5 is required for cooperative filament assembly but not involved in RNA binding [10,26,27].The CARDs of MDA5 nucleate the assembly of MAVS into its active polymeric form, which can be promoted by K63-linked polyubiquitin chains [10,28,29]. LGP2 has similar helicase and CTD domains and recognize the dsRNA with the similar protein-RNA contacts as those of RIG-I and MDA5. While LGP2 lacks the tandem CARDs [30] and does not recruit MAVS or induce MAVS signalling [1]. Evaluating the domain architecture of all the putative cgRLRs, except those seemly are incomplete duplicates, the protein EKC38303.1 is unique that it contains only one N-terminal CARD. Whether or not the single CARD is sufficient to initiate downstream signaling is yet to be investigated further. Another protein, EKC24604.1, is unique in that it's a combination of a helicase domain and a CLECT (C-type lectin /C-type lectin-like) domain. Animal C-type lectins are involved in functions such as extracellular matrix organization, endocytosis, complement activation, pathogen recognition, and cell-cell interactions. It is possible that this kind of combination carry out the similar function as RLRs do in mammals.

In conclusion, it is surprising that so many putative cgRLRs with a variety of domain architectures encoded in the genome of *C. gigas*. The study of invertebrate immunity can give us some insight into the evolution of human innate immunity [31]. Elucidation of the putative function of RLRs in antivirus response of invertebrates may help to design novel antiviral strategies for virus control in Oyster which is an important economic species. It is valuable to research the functions of cgRLRs and their activities in oyster antiviral immunity.

References

[1] Reikine S, Nguyen J B and Modis Y 2014 Pattern recognition and signaling mechanisms of RIG-I and MDA5 *Front. Immunol.* **5** 342
[2] Gack M U 2014 Mechanisms of RIG-I-like receptor activation and manipulation by viral pathogens *J. Virol.* **88** 5213-6
[3] Rodriguez K R, Bruns A M and Horvath C M 2014 MDA5 and LGP2: Accomplices and antagonists of antiviral signal transduction *J. Virol.* **88** 8194-200
[4] Yoo J S, Kato H and Fujita T 2014 Sensing viral invasion by RIG-I like receptors *Curr. Opin. Microbiol.* **20** 131-8
[5] Fitzgerald M E, Rawling D C, Vela A and Pyle A M 2014 An evolving arsenal: Viral RNA detection by RIG-I-like receptors *Curr. Opin. Microbiol.* **20** 76-81
[6] Rawling D C and Pyle A M 2014 Parts, assembly and operation of the RIG-I family of motors *Curr. Opin. Struct Biol.* **25** 25-33

[7] Goubau D, *et al* 2014 Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates *Nature* **514** 372
[8] Kowalinski E, Lunardi T, McCarthy A A, Louber J, Brunel J, Grigorov B, Gerlier D and Cusack S 2011 Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA *Cell*. **147** 423-35
[9] Peisley A, Lin C, Wu B, Orme-Johnson M, Liu M, Walz T and Hur S 2011 Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition *Proc. Natl. Acad. Sci.USA*. **108** 21010-5
[10] Berke I C and Modis Y 2012 MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA *EMBO. J*. **31** 1714-26
[11] Talmage S C and Gobler C J 2010 Effects of past, present, and future ocean carbon dioxide concentrations on the growth and survival of larval shellfish *Proc. Natl. Acad. Sci. USA.* **107** 17246-51
[12] Walters E T and Moroz L L 2009 Molluscan memory of injury: evolutionary insights into chronic pain and neurological disorders *Brain. Behav. Evol.* **74** 206-18
[13] Zhang G, *et al* 2012 The oyster genome reveals stress adaptation and complexity of shell formation *Nature* **490** 49-54.
[14] Altschul S F, Gish W, Miller W, Myers E W and Lipman D J 1990 Basic local alignment search tool *J. Mol. Biol*. **215** 403-10
[15] Marchler-Bauer A, *et al* 2011 *CDD:* a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res*. **39** (Database issue) 225-9
[16] Edgar R C 2004 MUSCLE: A multiple sequence alignment method with reduced time and space complexity BMC. *Bioinformatics* **5** 113
[17] Tamura K, Stecher G, Peterson D, Filipski A and Kumar S 2013 MEGA6: Molecular evolutionary genetics analysis version 6.0 *Mol. Biol. Evol*. **30** 2725-9
[18] Langevin C, Aleksejeva E, Passoni G, Palha N, Levraud J P and Boudinot P 2013 The antiviral innate immune response in fish: evolution and conservation of the IFN system *J. Mol. Biol.* **425** 4904-20
[19] Pietretti D and Wiegertjes G F 2014 Ligand specificities of Toll-like receptors in fish: Indications from infection studies *Dev. Comp. Immunol*. **43** 205-22
[20] Wang P H, Weng S P and He J G 2014 Nucleic acid-induced antiviral immunity in invertebrates: An evolutionary perspective *Dev. Comp. Immunol.* **48** 291
[21] Green T J and Montagnani C 2013 Poly I: C induces a protective antiviral immune response in the Pacific oyster (Crassostrea gigas) against subsequent challenge with Ostreid herpesvirus (OsHV-1 muvar) *Fish. Shellfish. Immun*. **35** 382-8
[22] Canesi L, Betti M, Ciacci C, Citterio B, Pruzzo C and Gallo G 2003 Tyrosine kinase-mediated cell signalling in the activation of Mytilus hemocytes: Possible role of STAT-like proteins *Biol. Cell*. **95** 603-13
[23] Zhang Y, Yu F, Li J, Tong Y, Zhang Y and Yu Z 2014 The first invertebrate RIG-I-like receptor (RLR) homolog gene in the pacific oyster Crassostrea gigas *Fish. Shellfish. Immun.* **40** 466-71
[24] Zeng W, Sun L, Jiang X, Chen X, Hou F, Adhikari A, Xu M and Chen Z J 2010 Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity *Cell*. **141** 315-30
[25] Berke I C, Yu X, Modis Y and Egelman E H 2012 MDA5 assembles into a polar helical filament on dsRNA *Proc. Natl. Acad. Sci. USA*. **109** 18437-41
[26] Li X, Lu C, Stewart M, Xu H, Strong R K, Igumenova T and Li P 2009 Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5. *Arch. Biochem. Biophys*. **488** 23-33
[27] Takahasi K, *et al* 2009 Solution structures of cytosolic RNA sensor MDA5 and LGP2 C-terminal domains: identification of the RNA recognition loop in RIG-I-like receptors *J. Biol.*

Chem. **284** 17465-74
[28] Jiang X, Kinch L N, Brautigam C A, Chen X, Du F, Grishin N V and Chen Z J 2012 Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response *Immunity*. **36** 959-73
[29] Hou F, Sun L, Zheng H, Skaug B, Jiang Q X and Chen Z J 2011 MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response *Cell*. **146** 448-61
[30] Yoneyama M, *et al* 2005 Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity *J. Immunol*. **175** 2851-8
[31] Ausubel F M 2005 Are innate immune signaling pathways in plants and animals conserved *Nat. Immunol*. **6** 973-9