Developmental and Comparative Immunology 51 (2015) 251–260

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Repertoire and evolution of TNF superfamily in *Crassostrea gigas*: Implications for expansion and diversification of this superfamily in Mollusca

Dahai Gao [a], Limei Qiu [a], Qiang Gao [a], Zhanhui Hou [a], Lingling Wang [a], Linsheng Song [b,*]

[a] Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, 7th Nanhai Road, Qingdao 266071, China
[b] Dalian Ocean University, Dalian 116023, China

ARTICLE INFO

Article history:
Received 19 January 2015
Revised 12 April 2015
Accepted 13 April 2015
Available online 21 April 2015

Keywords:
Crassostrea gigas
TNFSF gene
Phylogenetic analysis
Gene duplication
Diversification

ABSTRACT

Tumor necrosis factor superfamily (TNFSF) members represent a group of cytokines participating in diverse immunological, pathological and developmental pathways. However, compared with deuterostomia and cnidaia, the composition and evolution of TNF homologous in protostomia are still not well understood. In the present study, a total of 81 TNF superfamily (TNFSF) genes from 15 mollusk species, including 23 TNFSF genes from *Crassostrea gigas*, were surveyed by genome-wide bioinformatics analysis. The phylogenetic analysis showed that 14 out of 23 *C. gigas* TNFSF genes in five clades exhibited orthologous relationships with *Pinctada fucata* TNFSF genes. Moreover, there were 15 *C. gigas* TNFSF genes located in oyster-specific clusters, which were contributed by small-scaled tandem and/or segmental duplication events in oyster. By comparing the sequences of duplicated TNFSF pairs, exon loss and variant in exon/intron length were revealed as the major modes of divergence in gene structure. Most of the duplicated *C. gigas* TNFSF pairs were evolved under purifying selection with consistent tissue expression patterns, implying functional constraint shaped diversification. This study demonstrated the expansion and early divergence of TNF superfamily in *C. gigas*, which provides potential insight into revealing the evolution and function of this superfamily in mollusk.

1. Introduction

Tumor necrosis factor (TNF) superfamily refers to a large family of pleiotropic cytokines, which are involved in diverse biological pathways, such as cell proliferation, immune regulation, inflammation, cell death and apoptosis (Aggarwal, 2003; Croft, 2009; Locksley et al., 2001). The majority of TNF superfamily (TNFSF) members are type II membrane proteins with an intracellular N terminus and an extracellular C terminus (Igaki and Miura, 2014; Locksley et al., 2001; Wiens and Glenney, 2011). The C terminus is modestly conserved between TNFSF members (20–30%), which is defined as TNF homology domain (THD). The signature THD region is composed of 10 β-strands, which ultimately form a compact "jellyroll" sandwich structure (Naismith and Sprang, 1998). TNFSF members can bind to their cognate receptors (TNF receptors, TNFRs) by three THD regions to initiate various signaling transduction processes (Bodmer et al., 2002).

TNFSF members have been identified in a wide range of lineages of Eumetazoa, such as deuterostomia, protostomia and cnidaia, indicating that this superfamily has experienced extremely long evolutionary histories (Wiens and Glenney, 2011). In deuterostomia, the TNFSF repertoire has been extensively studied and putative full sets of 19 TNFSF members have been well characterized in human (Locksley et al., 2001), 18 members in zebrafish (Wiens and Glenney, 2011) and 24 members in amphibian (Huang et al., 2008). Furthermore, the evolutionary patterns of TNFSF members in deuterostomia have also been well investigated. Based on phylogenetic analyses, it is proposed that the majority of vertebrates TNFSF genes maintain orthologous relationships with human TNFSF genes, and TNFSF members suggesting have experienced conserved evolutionary conservation of TNFSF members since the appearance of vertebrates (Glenney and Wiens, 2007; Wiens and Glenney, 2011). The conserved relationships can even be traced to some basal deuterostomia groups; there are four TNFSF genes in *Strongylocentrotus purpuratus* (purple sea urchin) identified as potential orthologous to two human TNFSF members (Hibino et al., 2006). Meanwhile, the paralogous clusters are found to be of great contribution to the expansion and formation of lineage/species-specific TNFSF members. As an example, 11 of 19 human TNFSF members are physically located in the MHC

* Corresponding author. Dalian Ocean University, 52nd Heishijiao Street, Dalian 116023, China. Tel.: +86 0411 84762652; fax: +86 0411 84762652.
E-mail address: lshsong@dlou.edu.cn (L. Song).

http://dx.doi.org/10.1016/j.dci.2015.04.006
0145-305X/

regions on four chromosomes, created by small-scale or genome-wide duplication events (Collette et al., 2003).

In contrast, the knowledge about the composition and evolution of TNFSF members in protostomia is limited, partly due to less identified TNFSF members. For example, in the model ecdysozoan species, there are only one and two TNFSF homologues identified in *Drosophila melanogaster* and *Bombyx mori* respectively (Christophides et al., 2002; Igaki et al., 2002). And only scattered TNFSF homologous genes have been characterized in crustacean species, such as *MjTNF* from kuruma shrimp (Mekata et al., 2010) and *LvTNFSF* from whiteleg shrimp (Wang et al., 2012). In mollusca lineage, three TNF homologous genes have been reported, which are *AbTNF-α* and *AbFas* from disk abalone (De Zoysa et al., 2009a, 2009b) and *CgTNF-1* from Pacific oyster (Sun et al., 2014). Recently, three molluscan genomes of *Crassostrea gigas*, *Pinctada fucata* and *Lottia gigantean* have been published (Simakov et al., 2013; Takeuchi et al., 2012; Zhang et al., 2012), providing good opportunities to study the composition and evolution of mollusk TNFSF members.

Economically, Pacific oyster *C. gigas* is one of the most important aquaculture mollusks. Because of the ability of its adaptation to diverse biotic and abiotic stress conditions, Pacific oyster has also become a suitable model organism for studying immune and stress response mechanisms (Zhang et al., 2014). The investigations of TNFSF members in *C. gigas* will provide a comprehensive scenario to better understand the evolution and function of this immune-related gene family in mollusks. The objectives of this study were to (1) survey the TNFSF repertoire by screening the publicly available molluscan genome and EST sequence databases, (2) reveal the evolutionary properties of the TNFSF members identified in *C. gigas* by comprehensive phylogenetic and gene duplication analysis, and (3) elucidate the conservation and divergence of duplicated *C. gigas* TNFSF members by investigating domain/gene organization, evolutionary rate and expression profiles.

2. Materials and methods

2.1. The searching and identification of mollusk TNFSF members

The molluscan genome sequences of *C. gigas*, *P. fucata* and *L. gigantea* were downloaded from OysterDB (http://oysterdb.cn/home.html), OIST Marine Genomics Unit (http://marinegenomics.oist.jp/genomes/gallery?project_id=11), and JGI (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html), respectively. TNFSF sequences were retrieved by Hidden Markov Model (HMM) search against protein database for each genome. To retrieve possible TNFSF members which have not been annotated, TBLASTN searches against each assembled genome were carried out. The novel loci found by TBLASTN method were manually annotated by FGENESH + program (http://linux1.softberry.com/berry.phtml?topic=fgenes_plus&group=programs&subgroup=gfs) (see Supplementary file S1, Fig. S1 for detailed pipeline). A total of 1,125,736 expressed sequence tag (EST) sequences of mollusk animals were downloaded from NCBI database on March 27, 2013, and all putative TNFSF sequences were identified by TBLASTN searches. Another 14 reported molluscan TNFSF members in literatures (De Zoysa et al., 2009a; Igaki et al., 2002; Philipp et al., 2012) and one unpublished TNFSF were also included in this study.

The domain organization analysis was performed to verify the putative TNFSF members retrieved from *C. gigas*, *P. fucata* and *L. gigantea* genome. The presence of THD domain (PF00229) was validated by SMART (http://smart.embl-heidelberg.de/) and Pfam (http://pfam.sanger.ac.uk/). And the presence/absence of transmembrance (TM) region was evaluated by TMHMM (http://www.cbs.dtu.dk/services/TMHMM/).

2.2. Alignment and phylogenetic analyses

Different strategies were employed to generate the alignment of the TNFSF proteins from 15 mollusks. The CLUSTALX 2.0 (Larkin et al., 2007) program was firstly applied, but the alignment was not very accurate in THD domains because of the uncertainty in sequence comparison. The HMMalign program of the HMMer software package (Eddy, 1998) was then used, and the alignment generated under this way was of high quality and the residues within THD domain regions were reasonably aligned. The alignment was manually edited, and the unreasonable information sites were trimmed (see Supplementary file S2 for aligned matrix).

The phylogenetic tree reconstructions were carried out using both the neighbor-joining (NJ) and maximum likelihood (ML) methods. For both NJ and ML analyses, robustness of the obtained tree topologies was assessed with 1,000 bootstrap replicates. NJ trees were calculated with MEGA 5.01 program (Tamura et al., 2011) using the parameters of p-distance correction and pairwise deletion of gaps. Whereas ML trees were obtained with PhyML (version 3.0) using the model of WAG with a discrete gamma distribution (Guindon et al., 2010), which were recommended by Prottest (Darriba et al., 2011).

2.3. Sequence analysis

Tandem duplications were recognized by a close phylogenetic relationship analysis among tandem arrayed TNFSF genes located on the same scaffold. Segmental duplications were recognized by micro-synteny analysis for the flanking regions of the oyster TNFSF gene pairs. A 50 kb region covering both upstream and downstream of each gene was included and compared using the DotPlot function of the PipMaker program (Elnitski et al., 2003).

For selection pattern analysis, a PAL2NAL program was used to generate the codon alignment based on the sequence alignment of corresponding protein (Suyama et al., 2006). Nonsynonymous (*Ka*) and synonymous (*Ks*) substitution rates of coding sequences for duplicate pairs were calculated using the KaKs_Calculator (Wang et al., 2010) with the Goldman–Yang (GY) algorithm (Goldman and Yang, 1994) and modified Yang–Nielsen algorithm (MYN) (Zhang et al., 2006), respectively. For sliding window method, the window size was set as 20 amino acids and the step size was set as 10 amino acids.

2.4. Expression analysis

The available tissue expression values for *C. gigas* TNFSf genes were retrieved from the transcriptome data released by oyster genome project (http://oysterdb.cn/home.html, Zhang et al., 2012). The RPKM values (Reads per Kilobase of exon Model per million mapped reads) were used to indicate the expression levels. The expression pattern was visualized through heat map, which was generated by GenePattern service (http://genepattern.broadinstitute.org/). The coefficiencies of correlation for duplicate genes pairs were assessed by Pearson's R value.

3. Results

3.1. The identified TNFSF homologous in Mollusca

A total of 81 molluscan TNFSF homologous were collected in this study (Table 1), including 48 non-redundant TNFSF members thoroughly retrieved from the genomes of Pacific oyster, pearl oyster and owl limpet; 19 sequences with THD domains identified from molluscan ESTs; and 14 collected molluscan TNFSF members.

These TNFSF members are distributed in three major classes of Mollusca, which are Bivalvia, Gastropoda and Cephalopoda (Table 1).