ISJ 10: 29-37, 2013 ISSN 1824-307X

RESEARCH REPORT

# A novel third complement component *C3* gene of *Ciona intestinalis* expressed in the endoderm at the early developmental stages

**T Hibino[1], M Nonaka[2]**

[1]*Faculty of Education, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City 338-8570, Japan*
[2]*Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

*Accepted March 29, 2013*

## Abstract

The third complement component (C3) in ascidian was reported to function as an opsonin to enhance phagocytosis and as a chemotactic factor for phagocytes, indicating that ascidian C3 works in mesodermal cavity as a humoral factor like vertebrate C3s. In the basal Eumetazoa, Cnidaria lacking mesodermal tissues, C3 was reported to work in an endodermal cavity. Evolution of structure and function of C3 is still to be clarified. Here we report the identification of the third *C3* gene, *CiC3-3*, in the genome of an ascidian, *Ciona intestinalis*. Phylogenetic analysis using the entire amino acid sequences of Eumetazoan C3s indicated that CiC3-3 possess a closer relationship to vertebrate C3, C4 and C5 than other ascidian C3s. Although CiC3-3 retained the $\alpha$-$\beta$ processing site and 6 cysteine residues in the C3a region, it lacked the intra-molecular thioester bond and the catalytic histidine residue. Instead, CiC3-3 had a unique insertion of about 70 residues long Lys/Arg-rich sequence. *CiC3-3* was expressed highly in the embryonic stages, but little in the adult in contradistinction to *CiC3-1* and *CiC3-2*. The expression of *CiC3-3* in early embryonic stages was restricted to endoderm similar to cnidarian C3s. Thus, the ascidian complement system could represent a unique evolutionary stage sharing a primitive endodermal function with Cnidaria, and newly developed humoral function with vertebrates.

**Key Words**: thioester-containing protein (TEP); complement C3; innate immunity; immunogenetics; tunicate; chordate

## Introduction

The vertebrate complement system comprises more than 30 proteins present in serum or on cell surface, and plays a pivotal role in innate immunity. This system is triggered by three different activation pathways, the classical, alternative and lectin pathways. These three pathways merge at the proteolytic activation step of the complement component 3 (C3) into C3a and C3b. Upon proteolytic activation, C3 changes its conformation exposing the intra-chain thioester bond at the molecular surface. The exposed thioester bond of C3b reacts with surface molecules of invading microbes and makes a covalent bond, resulting in covalent tagging of microbes with C3b. Covalently attached C3b works as opsonin to induce phagocytosis, and also induces assembly of the terminal components of complement (TCCs: C6~C9) into a membrane-attack complex that can damage the membrane of certain pathogens (Murphy, 2011). The proteins possessing a similar domain structure as vertebrate TCCs are present in ascidian and amphioxus. However, these proteins may not be activated through the complement system of ascidian and amphioxus because they lack an essential domain for interaction with C5b (Nonaka and Kimura 2006).

The released smaller C3a fragment is an anaphylatoxin to induce inflammation. The C3 subfamily including C3, C4 and C5 is a member of thioester bond-containing protein (TEP) family, together with the non-complement TEP subfamilies such as the α2-macroglobulin (A2M) and CD109 subfamilies. The C3 subfamily members are distinguished from the A2M and CD109 subfamily members by the presence of the anaphylatoxin (ANA) and C-terminal of C3, C4 ,C5 (C345C) domains unique to the C3 subfamily (Sekiguchi *et al.,* 2012). Genes orthologous to vertebrate C3 have been identified not only from invertebrate deuterostome such as sea urchin (Al-sharif *et al.,*

*Corresponding author*:
Masaru Nonaka
Department of Biological Sciences
Graduate School of Science
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
E-mail: mnonaka@biol.s.u-tokyo.ac.jp

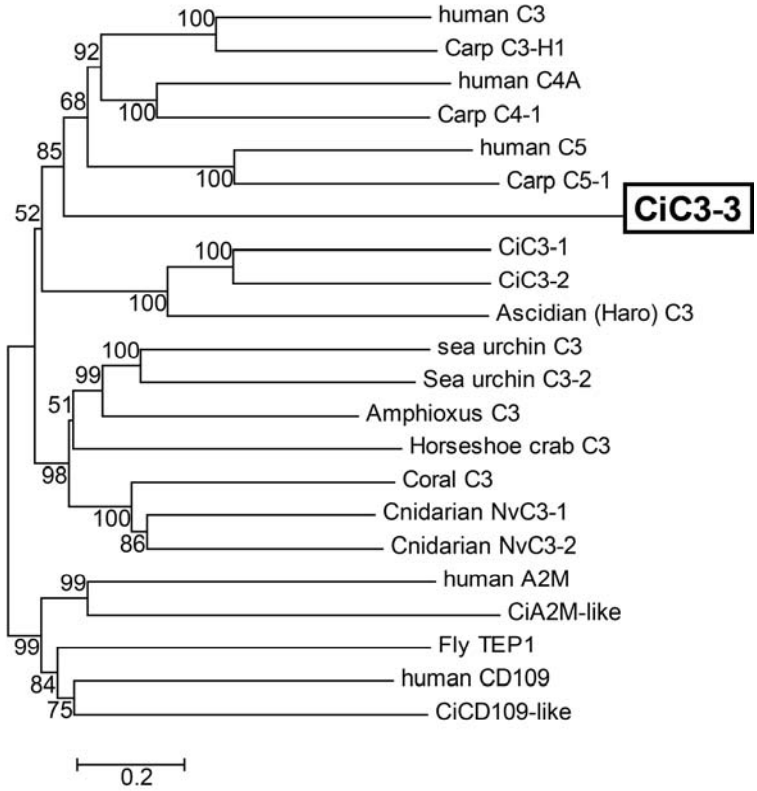


**Fig. 1** Phylogenetic tree of TEP family members constructed by the Neighbor-Joining method using the entire amino acid sequences. Bootstrap values higher than 50 % are indicated in the tree. Accession numbers of each entry are; human C3, C4A, C5, A2M, and CD109 (NP_000055, P0C0L4, AAA51925, P01023,and NP_598000), carp C3H1, C4-1, and C5-1 (BAA36619, BAB03284, and BAC23057), ascidian, *Halocynthia roretzi* C3 (BAA75069), *C. intestinalis* CiC3-1, CiC3-2, CiA2M-like, and CiCD109-like (NP_001027684, CAC85958, XP_002124325, and NP_001027688), sea urchin C3 and C3-2 (NP_999686 and Spbase: SPU_000997), horseshoecrab C3 (AAQ08323), amphioxus C3 (BAB47146), coral C3 (AAN86548), cnidarian NvC3-1 and -2 (AB450038 and AB450040), and fly TEP1 (NP_523578).

1998; Hibino *et al.,* 2006; Rast *et al.,* 2006) and amphioxus (Huang *et al.,* 2008), but also from protostomes such as horseshoe crab (Zhu *et al.,* 2005; Kawabata *et al.,* 2009) and spider (Sekiguchi *et al.* 2012) as well as cnidarian coral (Miller *et al.,* 2007) and sea anemone (Kimura *et al.,* 2009). The presence of C3 in Cnidaria indicated that the C3 gene has been established prior to the divergence of cnidarian from bilaterian (Nonaka, 2011). In contrast to its wide distribution, C4 and C5 has only been identified in jawed vertebrate, suggesting that C4 and C5 were derived from a C3-like common ancestor by gene duplication in the early stage of jawed vertebrate evolution (Nonaka and Takahashi, 1992).

A tunicate, *Ciona intestinalis* (Urochordata) has been an attractive research model for developmental biology for more than a century (Satoh *et al.,* 2003). The recent accumulation of genome-wide sequence information showed that not cephalochordate but urochordate is a sister group of vertebrate, indicating that *C. intestinalis* is one of the most important species for understanding the origin and evolution of vertebrates (Dehal *et al.,* 2002, Putnam *et al.,* 2008). The *C. intestinalis* genome analysis revealed that this animal possesses several genes for complement components: two C3s, three Bf/C2s, 10 of C6/C7/C8/C9/perforin and so on. An ancestor of the two C3-like genes seems to have diverged from a common ancestor of vertebrate C3/C4/C5 and has duplicated into two genes in the *Ciona* lineage (Azumi *et al.,* 2003a). Using *C. intestinalis*, in-depth expressed sequence tag and large-scale oligo-DNA microarray analyses have been advanced, which identified gene expression profile during the life cycle (Azumi *et al.,* 2003b; Satou *et al.,* 2002, 2003). Interestingly, *C3s, MASP, factor B (Bf), MBP* and two genes of complement *C6-like* were expressed only in the adult stages. On the other hand, *C1q-like* and two other genes of complement *C6-like* were expressed in the middle of the embryonic stages and maintained their expression level during the adult stages (Azumi *et al.,* 2007). The absence of C3 expression during the developmental stages could be explained by one of the following two hypotheses: (1) since *C. intestinalis* develops directly and metamorphosis in a day after fertilization, protection against infection which is considered to be the most

important physiological function of C3 is unnecessary in this short period, or (2) unidentified C3 is working under the developmental stage.

In this study, we report a novel *C. intestinalis C3* gene, *CiC3-3*, that belonged to a different clade from known ascidian *C3* genes in a phylogenetic tree, and that was specifically expressed in endoderm of the embryos.

## Materials and Methods

*Adults and embryos*

Adults of the ascidian *C. intestinalis* were provided from Misaki Marine Biological Station, the University of Tokyo through National Bio-Resource Project (NBRP) of MEXT, Japan. The adults were surgically dissected to draw eggs and sperm. Fertilized eggs were incubated at devitellination medium containing 0.065 % actinase E (Kaken Co. Ltd.) and 1.3 % sodium thioglycolate in sea water at a pH of approximately 10, to devitellinate chemically (Satou *et al.,* 2001). Devitellinated embryos were reared at 18 °C in agar-coated plastic dished filled with filtered sea water containing 50 μg/ml penicillin and 100 μg/ml streptomycin.

*Gene identification in* C. intestinalis *genome database*

Deduced amino acid sequences of all computationally predicted proteins were downloaded from the website, Ensembl *C. intestinalis* database (http://www.ensembl.org/) (Hubbard *et al.,* 2007), and the Ghost Database: *C. intestinalis* genomic and cDNA resources (http://ghost.zool.kyoto-u.ac.jp/indexr1.html) (Satou *et al.,* 2005). A typical C3 protein contains multiple domains; A2M_N, A2M_N2, A2M, ANATO, A2M_comp, A2M_recep, C345C, whose profile HMMs were downloaded from the Pfam website (http://pfam.sanger.ac.uk/) (Bateman *et al.,* 2004). HMMER (Eddy *et al*., 1998) was used to identify PFAM domain profile matches to *C. intestinalis* protein models. The deduced amino acid sequences of identified protein models were aligned with those of known *C. intestinalis* C3, CiC3-1 and CiC3-2 and Ciα2-macroblobulins to find out a novel C3 gene model.

*Cloning of a novel* C3 *gene of C. intestinalis*

*C. intestinalis* cDNA was synthesized from the adult tissues containing gills and blood cells, and used as a template for PCR amplification. To confirm the nucleotide sequences, especially exon-intron boundary, of the novel identified C3 (CiC3-3) gene model, RT-PCR was performed using primers that were designed at the ends of 5' UTR (forward 5'-TTGGAAAGCCGTACTATGCGACACG-3') and 3' UTR (reverse 5'-TGCTTTGGCAATATACACGTGGCAGT-3') of the gene model. Probably the nucleotide length of the predicted gene model of 5.7 kbp was too long for RT-PCR, the entire length of CiC3-3 could not be amplified by using these primers. We then designed other primers at the middle of the gene models (forward 5'-TGGAACAATCGCTGCTGCTGTAA-3', reverse 5'-ATGCCTTCTGGGACCACATTCAA-3'). ExTaq DNA polymerase (Takara) was used for this PCR. The cDNA fragments were cloned into pCR2.1-TOPO vector (Invitrogen) and were sequenced using vector specific primers or gene specific primers.

*Domain Prediction and phylogenetic analysis*

Domain structure of the CiC3-3 was predicted by the SMART program (http://smart.embl-heidelberg.de/) (Letunic *et al*., 2002). The e-value for the domain confidence was assessed by HMMER3 on the SMART program. Multiple alignment of the amino acid sequences among *C. intestinalis* and human C3s was done by ClustalW on the MEGA5 program (Tamura *et al*., 2011), as well as by eyes. Based on this alignment, phylogenetic trees were constructed using full-length amino acid sequence information or A2M_comp domain region that was extracted using by the SMART program. The neighbor-joining (NJ) method (Saitou and Nei 1987) using MEGA5 excluding gaps by pairwise deletion was performed. The reliability for internal branches was assessed by the 1000 bootstrap replications.

*Gene expression analysis using* Ciona *database and whole mount* in situ *hybridization*

The *C. intestinalis* protein database (CIPRO 2.5) integrates not only protein database, but also transcriptome database including large-scale EST analysis and DNA microarray data (Endo *et al*., 2011). We extracted the expression data of the three *CiC3* genes from the website and integrated the data in a graph. Whole mount *in situ* hybridization was performed based on the previously described protocol (Ogasawara *et al.,* 2001; Satou *et al*., 2001) with some modifications. For antisense or sense ribonucleotide probe for *CiC3-3,* 544 bp of the 3' end of coding region that covers the full length of the C345C domain and subsequent stop codon for *CiC3-3* was cloned into pTAC-2 with DynaExpress TA PCR Cloning Kit, and the probes were subsequently synthesized using Digoxigenin (DIG) RNA labeling mix and T7 or SP6 RNA polymerase (Roche). Embryos were fixed with 4 % paraformaldehyde in 0.1 M MOPS (pH 7.5), 0.5 M NaCl at 4 °C overnight. The developmental stages of fixed embryos were determined following Hotta *et al.* (2007). The fixed embryos were washed three times with PBST (phosphate-buffered saline containing 0.1 % Tween-20), then partially digested with 2 μg/ml proteinase K in PBST for 20 min at 37 °C. They were washed twice with PBST, subsequently post-fixed with 4 % paraformaldehyde in PBST for 1 h at room temperature (RT), and then washed three times with PBST. After prehybridization at 50 °C for 1 h, the embryos were hybridized at 50 °C for 24 h in the following buffer. The hybridization buffer contained 50 % formamide, 5x Denhardt's solution, 100 μg/ml yeast RNA, 0.1 % Tween-20, and 0.2 μg/ml DIG-labeled RNA probes. After hybridization, the embryos were washed three times with 2xSSC, 50 % formamide, 0.1 % Tween-20 at 50 °C for 15 min, then washed three times with 1xSSC, 50 % formamide, 0.1 % Tween-20 (A) at 50 °C for 15 min. Next they were washed twice with 1:1, (A): PBST at RT for 10min, and then washed three times with PBST at RT for 3 min. After the series of washing,

the specimens were blocked with 0.5 % blocking reagent (Roche) in PBST at RT for 30 min. They were immersed in 1/2000 Anti-DIG-AP fab fragments (Roche) diluted with PBST at RT for 6 h. The embryos were washed 4 times with PBST for 10 min and then washed twice with alkaline phosphatase buffer (0.1 M Tris-HCl (pH 9.5), 50 mM $MgCl_2$, 0.1 M NaCl) for 10 min. For signal detection, the embryos were incubated with NBT/BCIP in the alkaline phosphatase buffer at RT overnight. The stained embryos were dehydrated in a graded series of ethanol, and then cleared in a 1: 2 mixture of benzyl alcohol/benzyl benzoate.

## Results

*Identification of the third complement* C3 *gene in* C. intestinalis

To find gene candidates encoding multiple domains of typical thioester containing protein (TEP) superfamily from *C. intestinalis*, all of the deduced amino acid sequences of the Fgenesh gene models and the GENSCAN gene models that computationally predicted from the *C. intestinalis* genome were searched by local HMMER program using profile HMMs containing the A2M_N, A2M_N2, ANATO, A2M, A2M_comp and C345C domains. Out of five gene models extracted by this analysis, four gene models matched with the already reported TEP genes. The other gene model, Fgenesh76597 or GENSCAN101558, predicted a 3,759 bp open reading frame corresponding to a 1,253 amino acid sequence containing the A2M_N2, ANATO, A2M and C345C domains. The same gene was contained in the recently uploaded KH gene models (ver. 2012) in Ghost Database: *C. intestinalis* genomic and cDNA resources (http://ghost.zool.kyoto-u.ac.jp/indexr1.html) (Satou *et al.,* 2005). This gene model, KH.C12.243.v1.A.SL1-1, with a longer nucleotide sequence than that of the Fgenesh/GENSCAN model predicted a 1,873 amino acid sequence containing the additional A2M_N domain at its N-terminus. Based on these gene models, several primers were constructed at the end or at the middle of the sequences, and then a novel *C3* gene candidate was cloned and sequenced. The cloned nucleotide sequences had a size of 5,946 bp (1,873 amino acid residues) that matched 98.7 % (5666/5739) to that of the KH gene model. The novel and the third complement *C3* gene in *C. intestinalis* was designated as *CiC3-3*.

**Table 1**

*Phylogenetic analysis using the Neighbor-joining method*

The deduced amino acid sequence of the CiC3-3 gene was aligned with the known Eumetazoan TEP superfamily genes using ClustalW program (data not shown). The phylogenetic tree was constructed based on the entire amino acid sequences using the NJ method (Fig. 1). The phylogenetic tree showed the presence of three subfamilies, the C3, A2M and CD109 subfamilies, supported with bootstrap percentages of 52, 98 and 99 %, respectively. In the C3 subfamily, CiC3-3 was grouped with vertebrate C3/C4/C5 with a 85 % bootstrap percentage, but not with other ascidian C3 lineage including CiC3-1, CiC3-2 and *H. roretzi* C3. This result together with a long branch length of CiC3-3 indicates that CiC3-3 is a highly derivative ascidian C3, whose evolutionary origin is still to be clarified by analyzing other ascidian species.

*Structural features of CiC3-3*

To reveal whether CiC3-3 conserves the primary structure as well as domain structures of the vertebrate C3 subfamily, the deduced amino acid sequences of CiC3-3 were aligned with CiC3-1 and -2, and human C3, C4 and C5. The SMART domain search was also performed to find the multiple domains (Fig. 2).

The alignment and domain search showed that CiC3-3 possesses a signal peptide for secretion, the α/β processing site (RXXR) for dividing into two subunit chains, two Cys residues involved in an inter-chain disulfide linkage between the α and β chains and a possible activation cleavage site (TTR) by the C3 convertase (Fig. 1, Table 1), suggesting that CiC3-3 is processed into α- and β-chains held together with the inter-chain disulfide bond similar to mammal C3s. The C3a anaphylatoxin (ANA) region of CiC3-3 contained the six Cys residues conserved by most C3a analyzed thus far. Since CiC3-1 and -2 possess only four of them, CiC3-3 showed a higher conservation in the C3a region. However, CiC3-3 lacked the thioester site, GCGEQ, and the catalytic His residue for cleavage of thioester. Moreover CiC3-3 also lacked the two Pro residues on both sides of the thioester site which are conserved even in human C5 lacking the thioester site. These results suggest that the 3D structure around the thioester site is markedly modified in CiC3-3 (highlighted in yellow and blue in Fig. 2, summarized in Table 1).

| | α - β | ANA | Thioester | Catalytic His | α - γ |
|---|---|---|---|---|---|
| **CiC3-3** | RKKR | CC,C,C,CC | - | - | - |
| CiC3-1 | RKKR | C,C,C,C | GCGEQ | H | - |
| CiC3-2 | RNKR | C,C,CC,C | GCGEQ | H | - |
| human C3 | RRRR | CC,C,C,CC | GCGEQ | H | - |
| Human C4A | RKKR | CC,C,C,CC | GCGEQ | - | RRRR |
| Human C5 | RPRR | CC,C,C,CC | - | - | - |

**Fig. 2** Sequence comparison of CiC3-3, CiC3-1, CiC3-2 and the human C3, C4B, and C5. The multiple sequence alignment of CiC3-3 with human and other *C. intestinalis* C3s was performed with ClustalW. A2M_comp domain region is boxed. Proteolytic cleavage sites are shown in bold letters. Conserved Cys residues in the C3a anaphylatoxin region are marked (*). The inter-chain disulfide bridges between the α/β chains are shaded. The thioester sites, catalytic His sites and KR-rich insertion are also annotated in each colored box.

*Phylogenetic analysis of A2M_comp domain region to reveal independent loss of thioester site and catalytic His.*

To reveal whether loss of the thioester site and catalytic His occurred independently in CiC3-3 and vertebrate C5 or not, we reconstructed a phylogenetic tree with the deduced amino acid sequences based on the A2M_comp domain located at the C-terminal side of the thioester site. Although the size of the usual A2M_comp domain is approximately 260 amino acid residues long, this domain of CiC3-3 expands to 336 residues due to an insertion of approximately 70 amino acid residues highly enriched in Lys and Arg. A similar Lys/Arg rich insertion has already been reported from two cnidarian C3s, Nv3-1, Nv3-2, although the insertion of cnidarian C3 was observed at the different region, much more C-terminal side. Therefore, the insertion of the Lys/Arg rich sequence into C3 occurred at least twice independently during the eumetazoa evolution. The NJ tree constructed using the amino acid sequences of the A2M_comp domain showed the essentially the same topology as the tree based on the full length information described above, except that CiC3-3 is separated far from C3 family (Fig. 3). The long branch of CiC3-3 indicates that the primary structure of A2M_comp region of CiC3-3 is highly divergent. Overall, these results indicate that CiC3 possesses well conserved domain organization similar to vertebrate C3/C4/C5 except for the thioester site and the subsequent A2M_comp domain.

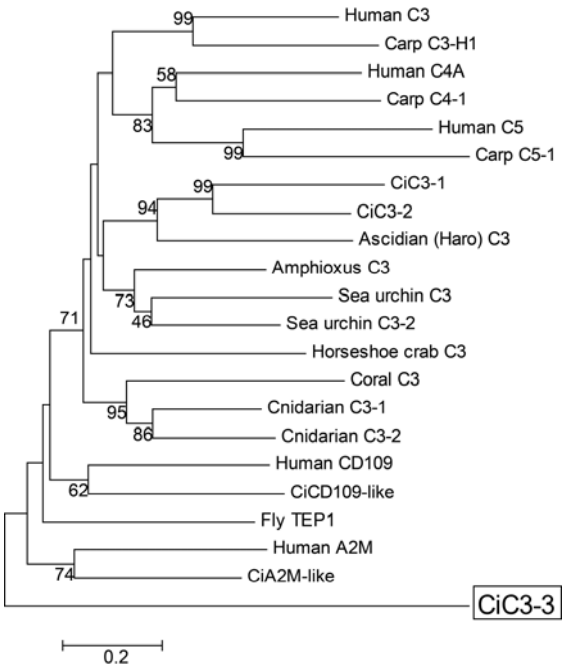


**Fig. 3** Phylogenetic tree of TEP family members constructed by the Neighbor-Joining method using the amino acid sequences of A2M_comp domain. Bootstrap values higher than 50 % are indicated in the tree. The genes in this tree are same as Figure 1.

*Spatial and temporal expression of the* CiC3-3 *gene*

To understand the gene expression pattern of *CiC3-3*, we first analyzed transcriptome data on the *C. intestinalis* protein database (CIPRO) (Endo *et al*., 2011), and compared gene expressions among *CiC3-1, CiC3-2* and *CiC3-3*. Both *CiC3-1* and *CiC3-2* were not expressed before the metamorphosis except for very slight expression in the tailbud stage, while both microarray and EST

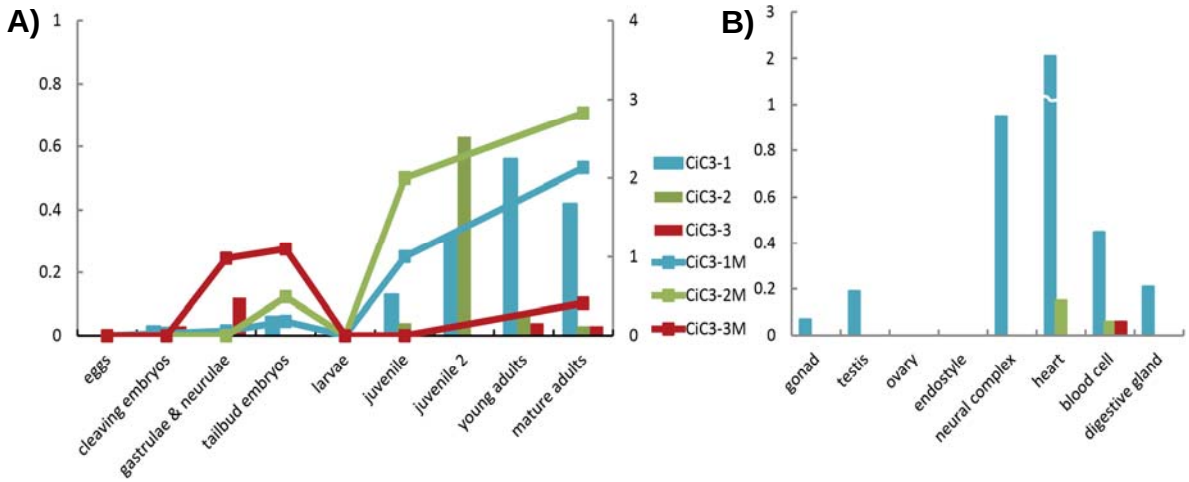


**Fig. 4** Comparison of expression intensities among *CiC3-1, CiC3-2* and *CiC3-3*. *CiC3-1, -2* and *-3* are shown in light blue, light green and red, respectively. The bars represent the EST data, while the lines represent the microarray data (labeled as CiC3-1M, -2M -3M). Y axes of Graphs A and B indicate relative expression levels. The results of the EST and microarray analyses are shown on the left and right sides, respectively. A: Expression profiles during the life cycle of *C. intestinalis.* The left side of the graph indicates the expression intensity for EST data, while the right side of the graph denotes the expression intensity for microarray data. B: Expression profiles of adult tissues. The bars denote expression level estimated by the EST analysis.

data showed that *CiC3-3* was significantly expressed from the gastrula to the tailbud stage (Fig. 4A). The expression of *CiC3-3* disappeared by the larva stage. After metamorphosis, *CiC3-1* and *CiC3-2* began to be expressed, whose intensities were getting stronger during maturation. In contrast to *CiC3-1* and *CiC3-2*, *CiC3-3* was not expressed by juvenile and was slightly expressed from the young adult to mature adult stages. The intensity of *CiC3-3* expression in mature adults is approximately 1/5 of *CiC3-1* and 1/7 of *CiC3-2* (Fig. 4A). The weak expression of *CiC3-3* was detected only in the blood cells. *CiC3-1* was ubiquitously expressed except for ovary and endostyle, and *CiC3-2* was expressed in heart and blood cell (Fig. 4B). These expression data indicate that *CiC3-3* is expressed in a contradistinctive manner from *CiC3-1* and *CiC3-2*.

To identify the spatial expression pattern of *CiC3-3* during the development of *C. intestinalis*, we next performed whole mount *in situ* hybridization using RNA probes of *CiC3-3*. *CiC3-3* began to be expressed in the invaginated cells of the early gastrulae (St. 11) (Fig. 5A). At the late gastrula stage (St. 13), almost all of the invaginated cells expressed *CiC3-3*. The *CiC3-3* expression was then restricted in the anterior end of the embryos (St. 14), especially the anterior ventral side of the invaginated cells strongly expressed *CiC3-3* (Figs 5D, E). The strong expression was observed in the endoderm of the trunk region, and weak expression was observed in the endoderm strand of the ventral midline of the tail region (St. 16 and 19) (Figs 5F, G, H). At the mid tailbud stage (St. 21) the *CiC3-3* expression was reduced and restricted only in the endoderm cells around the endodermal cavity (Figs 5I, J). These expression data indicates that *CiC3-3* is specifically expressed in the endoderm of embryos, and ceases its expression before hatching into the larvae.

**Discussion**

It had been reported that the number of complement *C3* gene is one in *H. roretzi*, and two in *C. intestinalis* (Nonaka *et al.,* 1999; Marino *et al.,* 2002; Azumi *et al.,* 2003). *H. roretzi* and *C. intestinalis* belong to the orders, Pleurogona and Enterogona, respectively, and are evolutionary far apart to each other (Turon *et al.,* 2004). The phylogenic analysis of C3 genes have indicated that the gene duplication event between CiC3-1 and CiC3-2 occurred in the Enterogona lineage after the divergence from Pleurogona (Marino *et al.,* 2002). The newly found CiC3-3 clustered with vertebrate C3, C4 and C5, rather than with CiC3-1, CiC3-2 and *H. roretzi* C3, although bootstrap percentage to support this clustering was not very high. This finding indicates the presence of two ancient C3 lineages in basal tunicates, the CiC3-1, CiC3-2 and *H. roretzi* C3 lineage and the CiC3-3 lineage. Two and three C3 genes were reported from the genomes of a sea urchin, *Strongylocentrotus purpuratus,* and an amphioxus, *Branchiostoma floridae*, respectively (Hibino *et al.,* 2006; Huang *et al.,* 2008). Thus all the basal deuterotomes whose genomes have been elucidated so far contain more than two C3 genes. Phylogenetic analysis showed that the multiple C3 genes of each species form species-specific cluster, indicating that gene

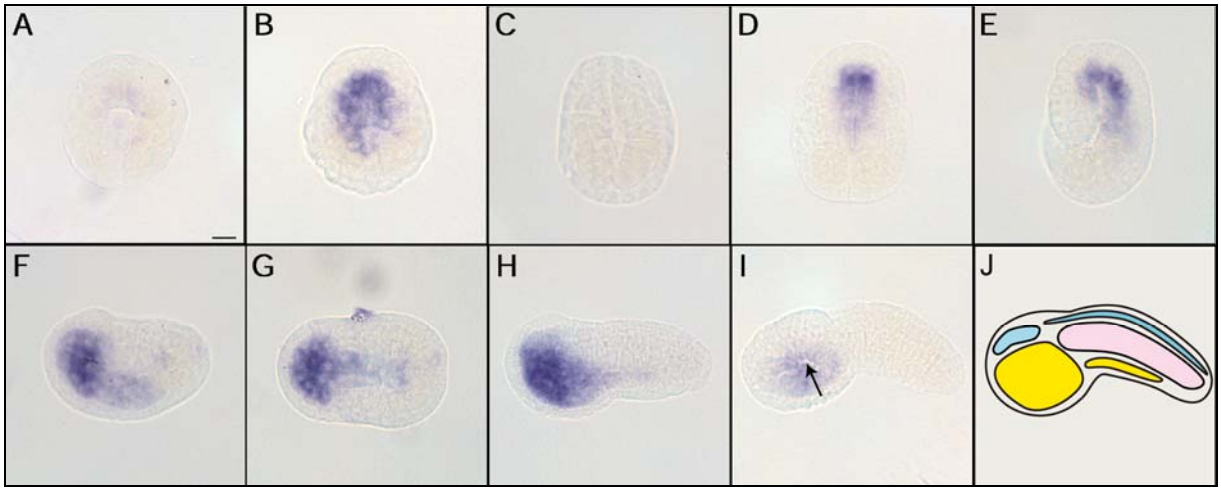


**Fig. 5** Spatial expression pattern of *CiC3-3* detected by whole mount *in situ* hybridization in the early gastrula through the mid tailbud stage embryos. The scale bar in A indicates 20 μm, and the magnification is the same for all pictures. A-E: anterior is toward to the top, F-J: anterior is toward to the left. F, H, I, J: ventral is toward to the bottom. A: vegetal view of early gastrula (4.9 hpf, St. 11), B-C: vegetal view of late gastrulae (5.9 hpf, St.13), C: no hybridization signal with sense probes, D: vegetal view of early neurulae (6.35 hpf, St. 14), E: lateral view of D, vegetal is toward to the right, F: lateral view of late neurulae (7.4 hpf, St. 16), G: ventral view of F, H: lateral view of early tailbud (9.3 hpf, St. 19), I, J: lateral view of mid tailbud (10 hpf, St. 21), arrow indicates endodermal cavity, J: Diagram of mid-tailbud corresponding to plate I. Yellow, light blue and pink denote endoderm, nervous system and notochord, respectively.

duplications occurred multiple times in each lineage. CiC3-3 is exceptional in this aspect, suggesting a unique evolutionary history of this gene.

CiC3-1 and CiC3-2 retain almost all domains and structural features of vertebrate C3, suggesting that they function as the central component of the ascidian complement system. Actually, the C3a fragment of CiC3-1 was demonstrated to induce chemotaxis of *C. intestinalis* hemocytes in the same way as vertebrate C3a (Pinto *et al.,* 2003). In contrast, CiC3-3 showed an unprecedentedly unique structure. First of all, CiC3-3 lacked the thioester site believed to be essential for covalent tagging of invading microorganisms by usual C3. Unlike vertebrate C5 which also lacks the thioester site but retains the basic residues of the thioester domain, CiC3-3 has a totally different sequence in this domain. Especially, the insertion of the highly Lys/Arg-rich sequence could have drastic structural and functional consequence since it provides extremely positive charge to this region. It is unlikely, therefore, that CiC3-3 play a similar function as vertebrate C3. However, the C3a region of CiC3-3 showed a higher conservation of Cys residues than those of CiC3-1 and CiC3-2, implicating in inflammatory process as anaphylatoxin.

In mammal, the C3 gene is mainly expressed in hepatocytes and macrophages (Lambris, 1988). In the ascidian *H. roretzi*, gastric caecum and blood cells have been identified as the sites of C3 gene expression (Nonaka *et al.,* 1999). The paraffin sections of the stomach in the adult of *C. intestinalis* have shown that both *CiC3-1* and *CiC3-2* are expressed only in the one type of blood cell, but not in the wall of the stomach (Marino *et al.,* 2002). Gene expression profile during the life cycle of *C. intestinalis* using the large-scale oligo-DNA microarray showed that not only CiC3-1 and CiC3-2, but also MASP, factor B, MBP and two genes of complement C6-like were expressed only in the adult stages (Azumi *et al.,* 2007). Two other genes of complement C6-like were expressed in the middle of the embryonic stages and maintained their expression level during the adult stages. In this study, *CiC3-3* showed a totally different temporal expression pattern during the life cycle from the other complement component genes of *C. intestinalis*. *CiC3-3* is prominently expressed during the embryonic stages when the other complement genes of *C. intestinalis* are hardly expressed. In adult stages, in contrast, *CiC3-3* is expressed at a very low level, whereas the other complement genes are expressed abundantly. Since interactions among components are essential for complement activation, these results suggest that CiC3-3 functions outside of the complement system.

Whole mount *in situ* hybridization revealed that *CiC3-3* was first expressed in the invaginating endoderm of the embryos. *C. intestinalis* develops in a direct developing manner, and the larvae do not undergo the differentiation of a functional gut. Thus, endodermal expression of *CiC3-3* does not necessarily indicate digestive function. When the gastrulation began, the expression started in the invaginated region, and it was continuously seen invaginating cells from the head endoderm through the endoderm strand to the ventral blastopore. After closure of the blastopore, it was strongly expressed around the endodermal cavity in the trunk. This expression pattern indicates the possibility that CiC3-3 is involved in development of certain embryonic region.

Complement C3 genes have been reported from basic metazoans, cnidarian coral, *Swiftia exserta* (Dishaw *et al.,* 2005), and cnidarian sea anemone, *Nematostella vectensis* (Kimura *et al.,* 2009). Another coral, *Acropora millepora C3* is expressed in undifferentiated endodermal cells of the embryos and larvae (Miller *et al.*, 2007), while *N. vectensis C3* is expressed in tentacles, pharynx, and mesentery in an endoderm-specific manner. Although all these cnidarian C3 possess the typical domain structure of vertebrate C3 unlike CiC3-3, a similar expression pattern during embryonic stages could imply that CiC3-3 and cnidarian C3 play some common developmental roles. If this is the actual case, cnidarian C3 has dual roles in development and immunity, which are divided into CiC3-3 and CiC3-1, 2, respectively in ascidians.

**References**


Al-Sharif WZ, Sunyer JO, Lambris JD, Smith LC. Sea urchin coelomocytes specifically express a homologue of the complement component C3. J. Immunol.160: 2983-2997, 1998.

Azumi K, De Santis R, De Tomaso A, Rigoutsos I, Yoshizaki F, Pinto MR, *et al.* Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: "waiting for Godot". Immunogenetics 55: 570-581, 2003a.

Azumi K, Takahashi H, Miki Y, Fujie M, Usami T, Ishikawa H, *et al.* Construction of a cDNA microarray derived from the ascidian *Ciona intestinalis*. Zool. Sci. 20: 1223-1229, 2003b.

Azumi K, Sabau SV, Fujie M, Usami T, Koyanagi R, Kawashima T, *et al.* Gene expression profile during the life cycle of the urochordate *Ciona intestinalis*. Dev. Biol. 308: 572-582, 2007.

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, *et al.* The Pfam protein families database. Nucleic Acids Res. 32: D138-D141, 2004.

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A,*et al.* The draft genome of Ciona intestinalis: Insights into chordate and vertebrate origins. Science 298: 2157-2167, 2002.

Dishaw LJ, Smith SL, Bigger CH. Characterization of a C3-like cDNA in a coral: phylogenetic implications. Immunogenetics 57: 535-548, 2005.

Eddy SR. Profile hidden Markov models. Bioinformatics 14: 755-763, 1998.

Endo T, Ueno K, Yonezawa K, Mineta K, Hotta K, Satou Y, *et al.* CIPRO 2.5: Ciona intestinalis protein database, a unique integrated repository of large-scale omics data, bioinformatic analyses and curated annotation, with user rating and reviewing functionality. Nucleic Acids Res. 39: D807-D814, 2011.

Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, *et al.* The immune gene repertoire encoded in the purple sea urchin genome. Dev. Biol. 300: 349-365, 2006.

Hotta K, Mitsuhara K, Takahashi H, Inaba K, Oka K, Gojobori T, *et al*. A web-based interactive developmental table for the ascidian Ciona intestinalis, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. Dev. Dyn. 236: 1790-1805, 2007.

Huang S, Yuan S, Guo L, Yu Y, Li J, Wu T, *et al.* Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Res. 18: 1112-1126, 2008.

Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, *et al.* Ensembl 2007. Nucleic Acids Res. 35: D610-D617, 2007.

Kawabata S, Koshiba T, Shibata T. The lipopolysaccharide-activated innate immune response network of the horseshoe crab. Inv. Surv. J. 6: 59-77, 2009.

Kimura A, Nonaka M. Genomic view of the evolution of the complement system. Immunogenetics 58: 701-713, 2006.

Kimura A, Sakaguchi E, Nonaka M. Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, *Nematostella vectensis*. Immunobiology 214: 165-178, 2009.

Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, *et al.* Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res. 30: 242-244, 2002.

Marino R, Kimura Y, De Santis R, Lambris JD, Pinto MR. Complement in urochordates: cloning and characterization of two C3-like genes in the ascidian *Ciona intestinalis*. Immunogenetics 53: 1055-1064, 2002.

Miller DJ, Hemmrich G, Ball EE, Hayward DC, Khalturin K, Funayama N, *et al*. The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss.Genome Biol. 8: 13, 2007.

Murphy K. Janeway's Immunology 8th Edition, Garland Science, 48-71, 2011.

Nonaka M. The complement C3 protein family in invertebrates. Inv. Surv. J. 8: 21-32, 2011.

Nonaka M, Azumi K, Ji X, Namikawa-Yamada C, Sasaki M, Saiga H, *et al.* Opsonic complement component C3 in the solitary ascidian, Halocynthia roretzi. J. Immunol.162: 387-391, 1999.

Nonaka M, Takahashi M. Complete complementary DNA sequence of the third component of complement of lamprey. Implication for the evolution of thioester containing proteins. J. Immunol.148: 3290-3295, 1992.

Ogasawara M, Minokawa T, Sasakura Y, Nishida H, Makabe KW. A large-scale whole-mount in situ hybridization system: Rapid one-tube preparation of DIG-labeled RNA probes and high throughput hybridization using 96-well silent screen plates. Zool. Sci. 18: 187-193, 2001.

Pinto MR, Chinnici CM, Kimura Y, Melillo D, Marino R, Spruce LA, *et al.* CiC3-1a-mediated chemotaxis in the deuterostome invertebrate *Ciona intestinalis* (Urochordata). J. Immunol.171: 5521-5528, 2003.

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, *et al.*The amphioxus genome and the evolution of the chordate karyotype. Nature 453: 1064-71. 2008.

Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW. Review - Genomic insights into the immune system of the sea urchin. Science 314: 952-956, 2006.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425, 1987.

Satoh N, Satou Y, Davidson B, Levine M. *Ciona intestinalis*: an emerging model for whole-genome analyses. Trends Genet. 19: 376-381, 2003.

Satou Y, Takatori N, Yamada L, Mochizuki Y, Hamaguchi M, Ishikawa H,*et al.* Gene expression profiles in *Ciona intestinalis* tailbud embryos. Development 128: 2893-2904, 2001.

Satou Y, Takatori N, Fujiwara S, Nishikata T, Saiga H, Kusakabe T, *et al. Ciona intestinalis* cDNA projects: expressed sequence tag analyses and gene expression profiles during embryogenesis. Gene 287: 83-96, 2002.

Satou Y, Kawashima T, Kohara Y, Satoh N. Large scale EST analyses in Ciona intestinalis - Its application as Northern blot analyses. Dev. Genes Evol. 213: 314-318, 2003.

Satou Y, Kawashima T, Shoguchi E, Nakayama A, Satoh N. An integrated database of the ascidian, *Ciona intestinalis*: Towards functional genomics. Zool. Sci. 22: 837-843, 2005.

Sekiguchi R, Fujito NT, Nonaka M. Evolution of the thioester-containing proteins (TEPs) of the arthropoda, revealed by molecular cloning of TEP genes from a spider, Hasarius adansoni. Dev. Comp. Immunol. 36: 483-489, 2012.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol.28: 2731-2739, 2011.

Turon X, Lopez-Legentil S. Ascidian molecular phylogeny inferred from mtDNA data with emphasis on the Aplousobranchiata. Mol. Phyl. Evol. 33: 309-320, 2004.

Zhu Y, Thangamani S, Ho B, Ding JL. The ancient origin of the complement system. EMBO J. 24: 382-394, 2005.