ISBN 978-979-792-665-6
International Proceeding

Phylogenetic and Structure Comparative Analysis of Complement Regulatory Membrane Proteins of Ginbuna Crucian Carp *Carassius auratus langsdorfii*

Indriyani Nur[1*], Hikari Harada[2], Ryota Nakamura[2], Masakazu Tsujikura[2], Tomonori Somamoto[2], Miki Nakao[2]

[1]Department of Aquaculture, Fisheries and Marine Science Faculty, Halu Oleo University Kendari, Indonesia 93232
∗E-mail: indri_noer@yahoo.com
[2]Laboratory of Marine Biochemistry, Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 812-8581, Japan

ABSTRACT

Regulation of the complement activation (RCA) has been shown to play a key role in controlling activation of the complement cascades as the main part of innate immune system. Humans RCAs have been studied most extensively and classified into soluble and membrane type. RCAs have also been partly characterized in a limited variety of non-mammalians species, including in bony fish, such as carp and zebrafish. In recent year, we have found three isoforms of complement regulatory membrane proteins (Tecrem) in ginbuna crucian carp at mRNA level, namely gTecrem-1, gTecrem-2, and gTecrem-3 in which structurally homologous to mammalian RCA-membrane type (CD46). In this study, phylogenetic analyses and comparative study on gTecrem structure were performed to clarify fundamental molecular information. Data based on mRNA level indicated that gTecrem structure principally conserved all the features of non-mammalians and human membrane-bound RCA. In addition, the findings showed that the number of Short Concensus Repeat (SCRs) in regulatory membrane-bound protein is different across isoforms and animal spesies.

Keywords: *complement activation, gTecrem, structure, phylogenetic*

INTRODUCTION

When faced with invading microbes, immune system has an arsenal of ways to fight off those pathogens. The complement has been traditionally believed as the major factor of the innate immune defense, as its name implies, with functions to contribute in enhancing antibacterial activity of humoral response. Indeed, its functions extend far beyond the elimination of microbes that are constantly present in the environment. It contributes substantially to various immune responses, regulating adaptive immunity, homeostasis, inflammatory, tissue regeneration, and lipid metabolism (Ricklin et al., 2010).

In the case of mammalian immune system, the complement is a highly complexes system consists of nearly 30 plasma and membrane-bound proteins which involved in its activation with at least three different proteolytic pathways to sense the pathogens infection. Virtually, complement activation results in the clearance pathogens, but when activated improperly, it may contribute to host tissue damage. A double edge sword then comes up due to dual roles of complement activation. Indeed, in humans, complement has been shown to associate with pathology induced by autoimmune reactions and chronic inflammation (Trouw and Daha, 2011). Therefore, regulation of the complement activation is of critical importance for homeostasis of the organism (Sjöberg et al., 2009).

To achieve normal level of complement activation, the mammalian host cell expresses several kinds of soluble and membrane-bound proteins that restrain complement activation by acting on several complement components. Among them, a group of regulatory protein acting on C4b and

C3b, termed regulators of complement activation (RCA), has been shown to play a key role in controlling activation of the complement cascades.

Still, RCA also has been identified in non-mammalian species such as chickens, frogs and lampreys (Inoue et al., 2001; Oshiumi et al., 2005, 2009; Kimura et al., 2004). Evidence of the existence of RCAs include in bony fish with a number of short consensus repeats (SCRs) as their peculiarities. Taken together, these data support the idea that SBP1 represents the complement-regulatory protein identified in barred sand bass and its homolog (SBCRP-1) (Dahmen et al., 1994; Zipfel et al., 1996), HCRF in the Japanese flounder (Katagiri et al., 1998), and the latest were ZRC1,2 in the zebrafish (Sun et al., 2010; Wu et al., 2012). Previously, we have found three isoforms of complement regulatory membrane proteins (Tecrem) in ginbuna crucian carp (*Carassius auratus langsdorfii*) at mRNA level, namely gTecrem-1, gTecrem-2, and gTecrem-3 in which structurally homologous to mammalian RCA-membrane type (CD46) (Nur *et al*., 2013).

In order to develop efficient strategy and approaches to control diseases for fish production, understanding of immune mechanism of fish is needed. Moreover, since most of fish have each characteristic of immune system and a variety of causes can influence their responses, thus the search for real molecular correlates of protection should be pursued with strong efforts. Therefore, this study was analyzed phylogenetic and comparative study on gTecrem structure to clarify fundamental molecular information and for further functional study.

MATERIALS AND METHODS

Cloning of Tecrem cDNA from ginbuna crucian carp. A gene-specific primer (5′-CAAGATGGGCTGTAAGGTGC-3′), corresponding to sequences containing the predicted 5′-untranslated region (UTR), was designed on the basis of the Tecrem sequence from common carp. First-strand cDNA from liver RNA was prepared using the SMART RACE cDNA amplification kit, according to the manufacturer's protocol.

3′-RACE PCR was performed under the following conditions: 40 cycles at 98 °C for 10 s, 61 °C for 10 s, and at 72 °C for 4 min. The amplified products were separated by agarose gel electrophoresis. The band of interest was extracted from the gel using FastGene Gel/PCR Extraction kit (Nippon Genetics Co., Tokyo, Japan), ligated to pGEM-T vector (Promega, Madison, USA), and introduced in DH5α competent cells. Plasmid was prepared by the alkaline mini-prep method (Birnboim and Doly, 1979).

Nucleotide sequencing. Nucleotide sequences were determined using the dideoxy chain termination method (Sanger et al., 1977) using a CEQ8800 DNA Analysis system and dye terminator cycle sequencing kit (Beckman Coulter Japan, Tokyo, Japan).

Computational sequence and phylogenetic analyses. Multiple sequence alignment of the amino acid sequences was performed using ClustalW (ver. 2) program available through internet (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Domain organization of gTecrem was predicted using SMART (http://smart.embl-heidelberg.de/) (Schultz et al., 1998). Signal peptide was predicted using SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei 1987) and displayed using MEGA4. Statistical significance of each branch was examined by the bootstrap percentages obtained from 1000 replications.

ISBN 978-979-792-665-6
International Proceeding

RESULTS AND DISCUSSION

Structurally, RCA molecules are largely or entirely composed of arrays of tandem globular modules termed SCRs, each of which consists of ~60 amino acid residues and characterized by a consensus sequence that includes four invariant cysteines, an almost invariant tryptophan and highly conserved distribution of prolines, glycines and hydrophobic residues (Seya, 1995; Kirkitadze and Barlow, 2001) (Fig. 1).

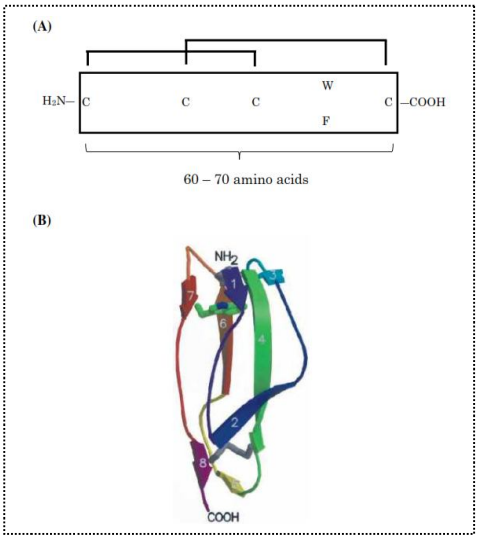


Figure 1. SCR domain structure. Primary structure of SCR domain (A) and tertiary structure generated by Molscript (Liszewski et al., 2000). β-Sheets are *numbered,* disulfide bonds at each terminus (between β-sheets 1 and 6; and 4 and 8) are shown. The highly conserved tryptophan (*green* ring) is shown with its *blue* nitrogen group.

Humans RCAs have been studied most extensively and classified into the following types: soluble type including factor H (FH), factor H-related proteins (FHR), and C4-binding protein (C4bp); membrane-bound type such as Complement Receptor type 1 (CR1, CD35), Complement Receptor type 2 (CR2, CD21), Decay-Accelerating Factor (DAF, CD55), and Membrane Cofactor Protein (MCP, CD46) (Liszewski et al., 1991; Liszewski et al., 2000; Jiang et al., 2001; Post et al., 1991) (Fig. 2).

In a phylogenetic point of view, RCAs have also been partly characterized in a limited variety of non-mammalians species, such as chickens (Inoue et al., 2001; Oshiumi et al., 2005; Oshiumi et al., 2009; Kimura et al., 2004) (Fig. 3). Some of them have been reported to show actual regulatory functions on the complement system, but there is no functional data to show any role of RCA in regulation of adaptive immune response of the lower vertebrates.

In bony fish, soluble RCA proteins similar to mammalian factor H have been identified from a few species. Namely, SBP1 and its homolog (SBCRP-1) have been identified from the barred sand bass (*Paralabrax nebulifer*) (Dahmen et al., 1994; Zipfel et al., 1996), and SBP1 showed a similar co-factor function as that of CD46. In the Japanese flounder (*Paralichthys olivaceus*), HCRF isoforms have been cloned but with no functional characterization (Katagiri et al., 1998). Zebrafish is a widely used model organism, whose complement system has been more extensively studied. In zebrafish, several factor H-like genes and another soluble form of RCA (named ZRC2) were found by in silico database mining. This study identified one complement factor H gene and four factor H-like genes in *Danio rerio*, and showed that they may originate by intra-chromosome duplication after the split of fish/mammalian common ancestor (Sun et al., 2010). In contrast, identification of membrane-bound RCA molecules

ISBN 978-979-792-665-6
International Proceeding

has been extremely limited, only reported for ZRC1 of zebrafish (Wu et al., 2012) (Fig. 4). Independently from the ZRC1 cloning, Tsujikura et al. (2015) has recently cloned cDNAs encoding membrane-bound RCA, designated teleost complement regulatory membrane protein (Tecrem) from carp (*Cyprinus carpio*) and zebrafish (*D. rerio*), and proven that carp Tecrem possesses a complement regulatory functions like CD46, suggesting that CD46 is a phylogenetically well conserved RCA molecule with ancient evolutionary origin.

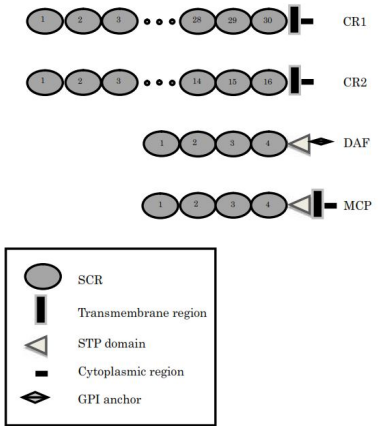


Figure 2. Primary structure of human membrane-bound RCA protein.CR1 contains 30 short consensus repeats (SCRs) while CR2 contain 16 SCRs. DAF (CD55) and MCP (CD46) contain 4 SCRs. Each SCR domain consists of approximately 60 amino acids with 4 conserved Cystein residues which form intradomain disulfide bonds. DAF attachs to the cell membrane via a glycosyl phosphatidylinositol (GPI)-anchor while MCP,CR1and CR2 insert into the cell membrane through transmembrane domains. DAF and MCP have Ser/Thr/Pro-rich (STP) regions between their SCR domains and the GPI-anchor or the transmembrane domain.

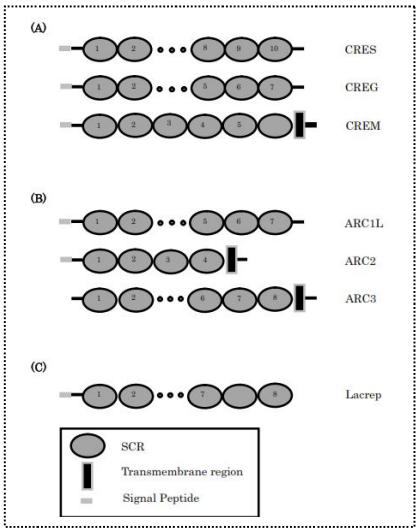


Figure 3. Domain structure of RCA protein in non-mammalian species. Chicken secretory RCA protein, CRES;chicken RCA with GPI-anchored protein, CREG; and chicken membrane RCA protein,CREM. Incomplete SCR is not numbered (A). Amphibian secretary RCA protein, ARC1; Amphibian membrane RCA protein, ARC2; Amphibian secretary RCA protein with a putative transmembrane region, ARC3 (B). Secretary RCA protein of Lamprey, Lacrep. SCR-like is not numbered.

Clonal triploid ginbuna crucian carp is a naturally occurring gynogenetic fish, and several clonal strains have been available for laboratory study (Nakanishi et al., 2011). These clonal fish strains can

ISBN 978-979-792-665-6
International Proceeding

serve as a model for studying innate and adaptive immune functions in both humoral and cellular aspects. In this reserach, complement component or RCA molecule has characterized in this species, hampering the molecular and cellular analyses on the relationship between complement in teleost. In the present study, we characterized the homologues of complement regulatory membrane protein (gTecrem) of all the SCRs in ginbuna crucian carp, then they were used to construct the phylogenetic tree showing the clustering relationships among the SCRs (Fig. 7).

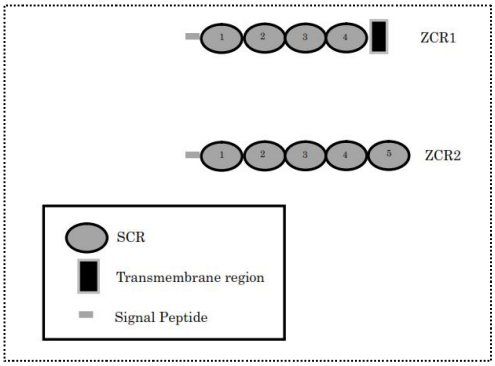


Figure 4. Primary structure RCA protein of zebrafish. Membrane-bound RCA, ZCR1, contain 4 SCRs with transmembrane (A) and soluble form RCA, ZCR2, contain 5 SCRs (B) were shown.

Interestingly, the three gTecrem isoforms consist of different number of SCR: gTecrem-1 has seven, whereas gTecrem-2 and gTecrem-3 have four SCRs (Fig. 5). These gTecrem isoforms are not the products of alternative splicing because their SCR domains have distinct sequences. Data sequences indicated that gTecrem structure principally conserved all the features of xenopus, chicken, and human membrane-bound RCA. The most significant difference among gTecrem isoforms is the number of SCRs, and the identities among the SCRs ranges from 54%–89% (Fig. 6). Also, data shows that gTecrem-1, gTecrem-2, and gTecrem-3 share a high degree of similarity to carp Tecrem and zebrafish Tecrem (up to 40%). It is therefore highly likely that Tecrem in ginbuna fish may be functional, also capable of regulating complement activation, especially binding to C3b.

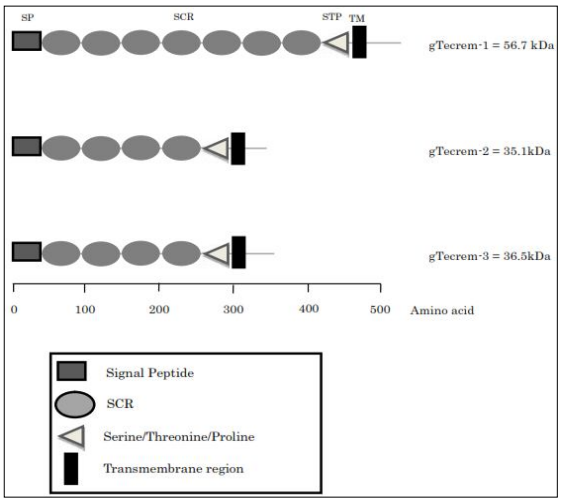
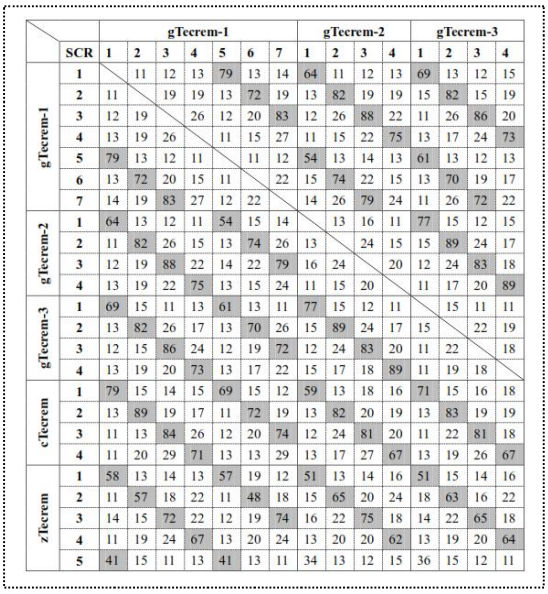


Figure 5. Domain organization of gTecrem isoforms. SP, signal peptide; SCR, short consensus repeat; STP, serine/proline/threonine-rich region; and TM, transmembrane region. Numbering of amino acids and molecular weights are shown.

ISBN 978-979-792-665-6
International Proceeding



		gTecrem-1							gTecrem-2				gTecrem-3			
	SCR	1	2	3	4	5	6	7	1	2	3	4	1	2	3	4
gTecrem-1	1		11	12	13	79	13	14	64	11	12	13	69	13	12	15
	2	11		19	19	13	72	19	13	82	19	19	15	82	15	19
	3	12	19		26	12	20	83	12	26	88	22	11	26	86	20
	4	13	19	26		11	15	27	11	15	22	75	13	17	24	73
	5	79	13	12	11		11	12	54	13	14	13	61	13	12	13
	6	13	72	20	15	11		22	15	74	22	15	13	70	19	17
	7	14	19	83	27	12	22		14	26	79	24	11	26	72	22
gTecrem-2	1	64	13	12	11	54	15	14		13	16	11	77	15	12	15
	2	11	82	26	15	13	74	26	13		24	15	15	89	24	17
	3	12	19	88	22	14	22	79	16	24		20	12	24	83	18
	4	13	19	22	75	13	15	24	11	15	20		11	17	20	89
gTecrem-3	1	69	15	11	13	61	13	11	77	15	12	11		15	11	11
	2	13	82	26	17	13	70	26	15	89	24	17	15		22	19
	3	12	15	86	24	12	19	72	12	24	83	20	11	22		18
	4	13	19	20	73	13	17	22	15	17	18	89	11	19	18	
cTecrem	1	79	15	14	15	69	15	12	59	13	18	16	71	15	16	18
	2	13	89	19	17	11	72	19	13	82	20	19	13	83	19	19
	3	11	13	84	26	12	20	74	12	24	81	20	11	22	81	18
	4	11	20	29	71	13	13	29	13	17	27	67	13	19	26	67
zTecrem	1	58	13	14	13	57	19	12	51	13	14	16	51	15	14	16
	2	11	57	18	22	11	48	18	15	65	20	24	18	63	16	22
	3	14	15	72	22	12	19	74	16	22	75	18	14	22	65	18
	4	11	19	24	67	13	20	24	13	20	20	62	13	19	20	64
	5	41	15	11	13	41	13	11	34	13	12	15	36	15	12	11

Figure 6. Amino acid sequence identity (%) of each SCR of ginbunaTecrem (gTecrem-1, gTecrem-2, and gTecrem-3), carp Tecrem, and zebrafishTecrem, numbers in the table are calculated using the program MEGA4. The identities higher than 40% are shaded to show the lineage of SCR modules.

The most different of RCA among species is the number of SCR (Fig. 2 – 4). The membrane-bound protein ZCR1 in zebrafish has five SCRs and molecular weight of 42.8 kDa (Wu et al., 2012). Human DAF and CD46 contain four SCRs, CR1 contains 30 SCRs, and mouse Crry contains five SCRs (Kim and Song, 2006; Liszewski et al., 1991). In lower vertebrates, xenopus ARC2 and ARC3 have four and eight SCRs, respectively, and carp and zebrafishTecrem contain four and five SCRs, respectively (Oshiumi et al., 2009). These findings indicated that the number of SCRs in regulatory membrane-bound protein is different across isoforms and animal spesies. Because the functional significance of the number of SCRs in complement regulatory proteins is still unclear, it would be interesting to understand the functional diversification among gTecrem isoforms.

Carp Tecrem possesses a complement regulatory functions like human CD46, suggesting that CD46 is a phylogenetically well conserved RCA molecule with ancient evolutionary origin (Tsujikura et al., 2015). Human RCA, excluding C4bpβ, fell into two distinct groups based on their sequence divergence, which were mapped on two separate loci of chromosome 1q32 (Krushkal et al. 2000). The phylogenetic tree constructed with neighbor-joining method showed that ginbuna Tecrem and zebrafish Tecrem clubbed together with carp Tecrem (Fig. 7). This indicated that Tecrem was member of RCA, derived from a common ancestor and similar function.

In summary, this study identifies three of complement regulatory membrane proteins (Tecrem) in ginbuna crucian carp. It is indicated that gTecrem structure principally conserved all the features of non-mammalians and human membrane-bound RCA, thus they may have similar function in immune system.

ISBN 978-979-792-665-6
International Proceeding

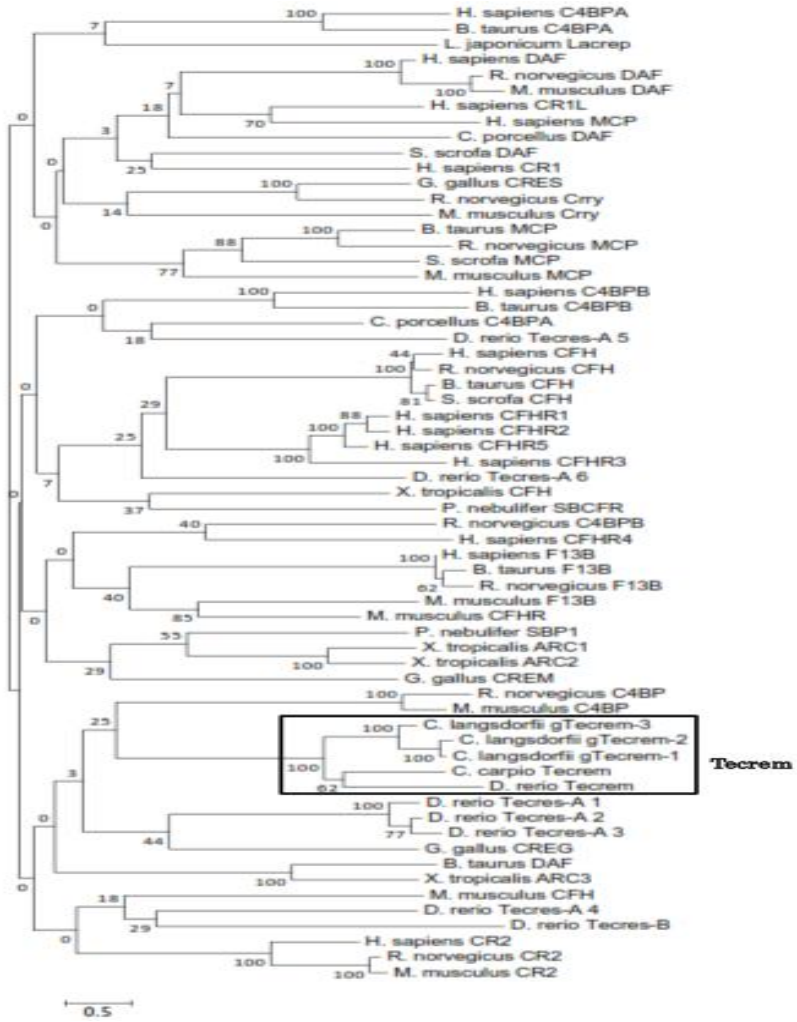


Figure 7. Phylogenetic tree of amino acid sequences of Tecrem.

The phylogenetic tree was constructed using neighbor joining (NJ) method and displayed using MEGA4. Bootstrap percentages are given. Amino acid sequences are obtained from the Gen Bank: Human (*H. sapiens*), cow (*B. taurus*), rat (*R. norvegicus*), guinea pig (*C. porcellus*), wild boar (*S. scrofa*), rodent (*M. musculus*), barred sand bass (*P. nebulifer*), frog (*X. tropicalis*), lamprey (*L. japonicum*), chicken (*G. gallus*), ginbuna crucian carp (*C. langsdorfii*), carp (*C. carpio*), and zebrafish (*D. rerio*). Abbreviations: C4BP, Complement 4-binding protein; DAF, Decay Accelerating Factor; F13B, Coagulation Factor XIII B polypeptide; CFHR, Factor H–related Protein; SBCFR, Sand Bass Cofactor Related Protein; SBP1, Sand Bass Cofactor Protein 1; Lacrep, Lamprey Complement Regulatory Protein; CRES, Chicken Complement Regulatory Secretory Protein; CREG, Chicken Complement Regulatory GPI-anchored Protein; CREM, Chicken Complement Regulatory Membrane-bound Protein; ARC, Amphibian Regulatory Complement Activation Protein; CFH, Complement Factor H; CR1, Complement Receptor 1; CR2, Complement Receptor 2;Crry, Rat Complement Regulatory Protein; MCP, Membrane Cofactor Protein; Tecres, Teleost Complement Regulatory Secretory Protein; Tecrem, Teleost Complement Regulatory Membrane-bound Protein

REFERENCES

Birnboim H.C. and Doly, J. (1979): A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7(6), pp: 1513-1523.

Dahmen A., Kaidoh T., Zipfel P.F., Gigli I. (1994): Cloning and characterization of a cDNA representing a putative complement-regulatory plasma protein from barred sand bass (*Parablax neblifer*). Biochem. J. 301(Pt 2), pp: 391–397.

Inoue N., Fukui A., Nomura M., Matsumoto M., Nishizawa Y., Toyoshima K., Seya T. (2001): A novel chicken membrane-associated complement regulatory protein: molecular cloning and functional characterization. J. Immunol.166(1), pp: 424–431.

Jiang H., Wagner E., Zhang H., Frank M.M. (2001): Complement inhibitor is a regulator of the Alternative Complement Pathway. J. Exp. Med. 194(11), pp: 1609–1616.

Katagiri T., Hirono I., Aoki T. (1998): Molecular Analysis of complement regulatory protein-like cDNA composed 12 tandem SCRs from Japanese flounder. Fish Pathol.33, pp: 351–355.

Kimura Y., Inoue N., Fukui A., Oshiumi H., Matsumoto M., Nonaka M.,Kuratani S.,Fujita T.,Masaru Nonaka N.,Seya T. (2004): A Short Consensus Repeat-Containing Complement Regulatory Protein of Lamprey That Participates in Cleavage of Lamprey Complement 3.The Journal of Immunol.

Kirkitadze M.D. and Barlow P.N. (2001): Structure and flexibility of the multiple domain proteins that regulate complement activation. Immunological Rev. 180, pp: 147-161.

Krushkal J., Bat O., Gigli I. (2000): Evolutionary relationships among proteins encoded by the regulator of complement activation gene cluster. Mol Biol Evol 17, pp: 1718–1730

Liszewski M.K., Post T.W., Atkinson J.P. (1991): Membrane cofactor protein (MCP or CD46): newest member of the regulators of complement activation gene cluster. Annu. Rev. Immunol. 9, pp: 431–455.

Liszewski M.K., Leung M., Cui W., Subramanian V.B., Parkinson J., Barlow P. N., Manchester M., Atkinson J.P. (2000): Dissecting Sites Important for Complement Regulatory Activity in Membrane Cofactor Protein (MCP; CD46).J. Biol. Chem. 275(48), pp: 37692-37701.

Nakanishi T., Toda H., Shibasaki Y., Somamoto T. (2011): Cytotoxic T cells in teleost fish.Dev. Comp. Immunol.35(12), pp: 1317–1323.

Nur, I., Harada, H., Tsujikura, M., Somamoto, T., Nakao, M. (2013): Molecular characterization and expression analysis of three membrane-bound complement regulatory protein isoforms in the ginbuna crucian carp *Carassius auratus langsdorfii*. Fish Shellfish Immunol. 35, pp: 1333-1337.

Oshiumi H., Shida K., Goitsuka R., Kimura Y., Katoh J., Ohba S., Tamaki Y., Hattori T., Yamada N., Inoue N., Matsumoto M., Mizuno S., Seya T. (2005): Regulator of complement activation (RCA) locus in chicken: identification of chicken RCA gene cluster and functional RCA protein. J.Immunol. 175(3), pp: 1724-1734.

Oshiumi H., Suzuki Y., Matsumoto Y., Seya T. (2009): Regulator of complement activation (RCA) gene cluster in *Xenopustropicalis*. Immunogenetics 61(5) , pp: 371–384.

Post T.W., Liszewski M.K., Adams E.M., Tedja I., Miller E.A., Atkinson J.P. (1991): Membrane cofactor protein of the complement system: alternative splicing of serine/threonine/proline-rich exons and cytoplasmic tails produces multiple isoforms that correlate with protein phenotype. J. Exp. Med. 174(1), pp: 93–102.

Ricklin D., Hajishengallis G., Yang K., Lambris J.D. (2010): Complement: a key system for immune surveillance and homeostasis. Nat.Immunol. 11(9), pp: 793-797.

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

Sanger F.S., Niclen S., Coulson A.R. (1977): DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U S A :74, pp: 5463–5467.

Schultz J., Milpetz F., Bork P., Ponting C.P. (1998): SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. U S A. 95, pp: 5857-5864.

Seya T. (1995): Human regulator of complement activation (RCA) gene family proteins and their relationship to microbial infection. Microbiol.Immunol. 39(5):295-305.

Sjöberg A.P., Trouw L.A., Blom A.M., (2009): Complement activation and inhibition: a delicate balance.Trends Immunol. 30(2), pp: 83-90.

Sun G., Li H., Wang Y., Zhang B., Zhang S. (2010): Zebrafish complement factor H and its related genes: identification, evolution, and expression. Funct.Integr.Genomics. 10(4), pp: 577–587.

Tsujikura, M., Nagasawa, T., Ichiki, S., Nakamura, R., Somamoto, T., Nakao, M. (2015): A CD46-like molecule functional in teleost fish represents an ancestral form of membrane-bound regulators of complement activation. J. Immunol. . 194, pp: 262-272.

Wu J., Li H., Zhang S. (2012): Regulator of complement activation (RCA) group 2 gene cluster in zebrafish: identification, expression, and evolution. Funct.Integr. Genomics 12(2), pp: 367–377.

Zipfel P.,F., Kemper C., Dahmen A., Gigli I. (1996): Cloning and recombinant expression of a barred sand bass (*Paralabraxnebulifer*) cDNA. The encoded protein displays structural homology and immunological crossreactivity to human complement/cofactor related plasma proteins. Dev. Comp. Immunol. 20(6), pp: 407–416.