

잉어톨루사접수체유전자 1(*tlr1*)의 배열과 진화

장성훈, 백은주

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《모든 과학자, 기술자들이 과학기술발전의 추세에 맞게 첨단과학과 기초과학발전에 힘을 넣어 나라의 과학기술을 세계적수준에 올려세우도록 하여야 합니다.》(《김정일선집》 중보판 제20권 62페이지)

포유동물에서 TLR(Toll-like receptor)족 유전자는 병원체결합분자형태(PAMPs)를 인식하고 염증성세포인자들과 I형인터페론을 발현시켜 병원성세균에 대한 유기체의 면역을 조절한다. 우리는 잉어를 연구대상으로 하여 물고기의 선천성면역에서 중요한 역할을 하는 *tlr1* 유전자의 배열과 진화적특성을 밝혔다.

재료와 방법

1) 실험에 리용한 프라이머

프라이머는 Primer premier 5.0프로그램으로 설계하였다.(표 1)

표 1. 잉어*tlr1* cDNA의 증폭에 리용한 프라이머

프라이머이름	배열(5'→3')	증폭위치
SMART II(TM) (정방향)	AAGCAGTGGTATCAACGCAGAGTACGCGGG	첫번째 사슬 cDNA의 증폭
CDSIII (반대방향)	AAGCAGTGGTATCAACGCAGAGTAC (T) 30V N	첫번째 사슬 cDNA의 증폭
gcT1-1F(정방향)	CGTTTAGAGAAGTGAGACAGC	열린읽기틀
gcT1-1R(반대방향)	ACCAGACTTTCCAGGTTTGT	열린읽기틀
gcT1-2F(정방향)	TTCTTCCTCAGATTCACCTCG	열린읽기틀
gcT1-2R(반대방향)	ACCAGACTTTCCAGGTTTGT	열린읽기틀
gcT1-3F(정방향)	TGCTGGTGGTTGATGTCTG	열린읽기틀
gcT1-3R(반대방향)	GGTAGTACTTGGAGGGAATGA	열린읽기틀
Tlr1F1(정방향)	CAAYCCHTTTCTMHWTGYACHTGTG	열린읽기틀
Tlr1R1(반대방향)	TCRTARTGRCACCACTCRCTCT	열린읽기틀
T1fR1(반대방향)	CATAAAGATGATGGGCGTTTGGGT	5'—말단
T1fR2(반대방향)	CGATCTTGCTAACAGTCTGAGGGC	5'—말단
T1G2(정방향)	CAAGCACACGCAATAACCTCCCAG	3'—말단
T1NG2(정방향)	CTCCCAGAAGAAGACTGAGGATATGG	3'—말단

2) cDNA의 전 배열의 증폭

SMART cDNA의 합성방법[1]을 리용하여 잉어의 두신으로부터 분리한 mRNA를 역번역시켜 주형cDNA를 얻는다. 다음 Vector NTI프로그램으로 다중배열상동성검색을 진행하고 진화적으로 보존된 TIR도메인영역배열을 찾아내어 축퇴프라이머(Tlr1F1/Tlr1R1)를 합성하

였다. 이 프라이머를 리용하여 PCR법으로 잉어 *tlr1* 유전자의 보존된 배렬을 증폭한다. 다음 증폭된 토막을 pMD18-T운반체(《Promega》)에 재조합하고 배렬분석을 진행한다. 배렬분석자료에 근거하여 5′-와 3′-말단을 증폭하기 위한 프라이머를 설계한다. SMART II와 T4G1(T4NG1), T4G2(T4NG2)와 CDS III 프라이머로 목적 유전자의 양쪽 말단을 증폭한다. 증폭된 5′- 및 3′-말단의 PCR산물들을 pMD18-T운반체에 재조합하고 배렬을 분석하였다.

3) 생명정보학적분석

tlr1 유전자의 진화분석은 ClustalW2 프로그램으로 진행하였다. 분석에는 줄말고기(*Danio rerio*, XP697531), 자지 보가지(*Takifugu. rubripes*, AC156430), 고릴라(*Gorilla gorilla*, BAG55017), 오랑우탄(*Pongo pygmaeus*, BAG55018), 긴팔원숭이(*Hylobates lar*, ACC68075), 붉은털원숭이(*Macaca mulatta*, NM001130424), 흰눈썹검은갯원숭이(*Cercocebus torquatus atys*, ABY64985), 메돼지(*Sus scrofa*, NP001026945), 소(*Bos indicus*, ABU86938), 양(*Ovis aries*, NM001135060), 흰생쥐(*Mus musculus*, NP109607), 닭(*Gallus gallus*, XP423937)의 *tlr1* 유전자배렬들이 리용되었다. 핵산배렬의 분석은 Vector NTI 프로그램으로 진행하였다. 단백질의 아미노산배렬과 분자량은 cDNA 암호배렬에 근거하여 BLAST 프로그램의 SwissProt에서 계산하였다.

결과 및 논의

1) 잉어 *tlr1* 유전자의 핵산 및 아미노산배렬에 대한 분석

SMART cDNA 합성방법을 리용하여 얻어낸 잉어 *tlr1* 유전자의 증폭산물들을 pMD18-T운반체에 클론화한 다음 배렬분석을 진행하였다.

그 결과에 의하면 5′-말단 및 3′-말단 증폭산물로서 1 300, 798bp 토막이 얻어졌으며 중간배렬로서 2 294bp의 배렬이 분석되었는데 이 배렬들은 호상 겹치므로 Vector NTI 프로그램으로 분석하면 2 948bp의 완성된 cDNA배렬(폴리A배렬을 제외)이 얻어졌다.(그림 1) 잉어 *tlr1* 유전자에서 5′- 및 3′-비번역배렬의 길이는 각각 267, 296bp이고 열린읽기틀의 길이는 2 385bp이며 3′-비번역배렬에 배렬의 불안정성을 보여주는 3개의 모티프(ATTTA)들이 있다. 이 배렬에 의하여 암호화되는 단백질의 크기는 794aa이다.

```

1      GAGCATATTCAACATTTAATTTCTTTTGCTTTCAAATGCTTCTGCGCTCACAGTAAAGTAGAGCAAGTGCAGAA
76     AAACCTAAATTTACATAGAAAAAACCCTAAAAAATGAAAAATACTGCTGCTATGCAACTACAATGCAAAAC
151    TAATGCTTGGGATTAAGTAAAAGTGAAAACTGTGAGAAGTCAAAAATGCGGTTGACTTGAAATGGGAAATGAAAT
226    CAAGAGTGTTTTTCTCTCTCAGATTCACCTCGTGTGAGAAAGATGAAGCCGTCATGCTGGTGGTTGATGTCTGTT
                                     M K P S C W W L M S V    11
301    TATGTCACGTGTTTACACCCATCTCTCGTCTTGCCATTAAAGGATCATAGTGAATTACTCATCTCAGAACCTG
      Y V T C L H P S L V L A I K R I I V N Y S S Q N L    36
376    TCCTCGGTTCTCCCGGCTGAAACCATCTACAGAGGACTTGGATCTGTGCTGAACCACATCCAGACGCTGAGC
      S S V P P G L K P S T E D L D L S L N H I Q T L S    61
451    GGCAAGACTTCAGCACTACACCACGACTCCATTTCTCTCAATCTGTCTGGAATATACTGGAACATAGACATG
      G K D F S T T P R L H F L N L S W N I L E N I D M    86
526    GACACATTTAACTCCACACCGCTTTGGAGATTTTGGACCTGTGCGACAAACAGACTCCAGAACCTCTCTGACCAA
      D T F N S T P A L E I L D L S H N R L Q N L S D Q    111
601    CCATACCTGCTTAAGGCGGGAATCTGCAGTCTCTAGACTTGTTCATCCAACATGTTTTCCGTCATGGCACTAGGG
      P Y L L K A G N L Q F L D L S S N M F S V M A L G    136
676    AAGGAATTTTCTACCTGAAACACCTGCAGTGGCTTGGCCTGAGTGCCATTTCATCGGCAACGAAGACTTTACT
      K E F S T T L K H L Q W L G L S A I S I G N E D F T    161
751    TACATTGCCAATCTTACTTTGAAAACGCTCTTATTAATACTGATAAGCTACAGAAGTATGAAGAGAGGAGTCTA
      Y I A N L T L K T L F I N T D K L Q K Y E E R S L    186
826    ATGGGTGCCCATGCAGAGAAGCCGTCATCGCCCTGTCCAACAAGACCTAGATATTGCAGTCGCTGATGATGTT
      M G A H A E K A V I A L S N K D L D I A V A D D V    211
    
```

901	TTTGCATCATTTAAAGAGGTAGAAATTCACCAAGATGGACAGCGAGATGAAGTTCTCCAGCAGATATGCCAGCGA	
	F A S F K E V E F T K M D S E M K V L Q Q I C Q R	236
976	GGAACAGTACGACCATTAGTCTAGAAATCTCCAAGTAAAAACCCTGGGAAGTTCTGACAGGCTGTGTAAAC	
	G T V R T I S L E I S K V K T T W E V L T G C V N	261
1051	ACAATACTGTCTATCTGCCATTCAACAGCTTTCCTTCTCAGACCTTACCATGACAAAAATGATAGACGGCAGGCCT	
	T I L S S A I Q Q L S F S D L T M T K M I D G R P	286
1126	GTGTCTCGAGCCGCTTGTGGAGTCCTTCTCTACAAGAAGAGCGTCTGTAACCGAATTTCATCTTCAACCAACA	
	V S S S R L L E S F S T R R A S V T E F I F N Q T	311
1201	GAGTCTATGACTTCTTCATAAGCATGCCAGCGAGAAAAGTTAGTTTAAACCAACGCCCATCATCTTTATGACC	
	E L Y D F F I S M P A R K V S L T Q T P I I F M T	336
1276	TGCCCTCAGACTGTTAGCAAGATCGAGGTGTGGACTTGTCCGATTGCGCCCTTACGAAAAACATCTTCTCGGTT	
	C P Q T V S K I E V L D L S D C A L T E N I F S V	361
1351	GGTCCAGATACAGAATGTAGTACCTAACAAACCTGGAAAGTCTGGTTCTGAGGGGCAACACCTGAAACATCTT	
	G P D T E C S T L T N L E S L V L R G N N L K H L	386
1426	GGACTGCTGACCTCCAGATCCATCTCATGAGCTCACTGAAATACGCTGACTTCAGTCAGAACCAAGCTTACTTAT	
	G L L T S R I H L M S S L K Y A D F S Q N K L T Y	411
1501	TCTGAGGAACAAGGCAAGTGCCTTGGCCTCGAAAGTTGCCACGTGGATTGTCTCAATGGGTTTGACCAG	
	S E E Q G K C V W P R K V A H V D L S S N G F D Q	436
1576	AGCATCTTCAAGTGTGTCGCCGATTCCATAAGAATCTTAAACCTCCGGAATAATCGAGTGACCACAGTCCCTGCA	
	S I F K C L P D S I R I L N L R N N R V T T V P A	461
1651	GATGTTGGGCTTTTGAAGACTTGAAGCATCTGGACCTAGCAGATAACCGCTTGTCTGGACCTTCCCACTTGTCAA	
	D V G V L E D L S I L D L A D N R L L D L P T C Q	486
1726	GCATTCCAAACCTGCAGAAGCTCTAGTAAGAACAACCTCCATCCATTGCGCTTTACCGGGAACCCCTGAGACT	
	A F P N L Q K L L V R T N S I H S P L P G T L E T	511
1801	TGTCCTCATCTCAAAGATCTCGACTTGAGCCGCAACCCCTTCATCTGCACTTGTGCTTTGCGAGAATTGCGAGCT	
	C P H L K D L D L S R N P F I C T C A L R E F A A	536
1876	CTCTCAAAAAACCAAGCAACACAACTCCTGGTGTGACTCTCGGCCACTCGCCGAAGGATATCAATGTAGTTAC	
	L I K N Q A T Q T P G V T L G H W P E G Y Q C S Y	561
1951	CCAGAATCCAGGAGCAACACCATGCTAAAAGACTTCTACCTTCTGAGATATCTGTGATGGCTGGACTCTGGCT	
	P E S R S N T M L K D F Y L P E I S C D G W I L A	586
2026	GTTACGATTCTGATCCCGACAATCACTCTGGTAGTTGCCATCAGCCTTCTTTGCCATCGGCTGGACGTGCCGTGG	
	V T I T L V V A I S L L C H R L D V P W	611
2101	TACCTCAAGATGATGTGGAATGGACACGCGCAAGCACCAGCAATAACCTCCAGAAAGACTGAGGATGATG	
	Y L K M M W K W T R A K H H A I T S Q K K T E D M	636
2176	GAGGGGTTGCGCTTTTCATGCATTCTGATCGTACAGCCAAAAGAACGAGACTGGGTCAAAGCTCAATTTCTTCCC	
	E G L R F H A F V S Y S Q K N A D W V K A Q F L P	661
2251	AAGCTCGAGGGCGATTGTGGGTGCGAGTTTGTCAACACGAGCGGGACTTTATCCCGGGAAGAGCATCGTTACAG	
	K L E G D C G L R V C H H E R D F I P G K T I V Q	686
2326	AACATCTCCGATGCATGAGCAGCAGCAGAAGTGCCTCTCGTGCTGTCTCATCTTTCATCCAGGCAATGG	
	N I L R C I E H S R R C V F V L S S H F I Q G S E W	711
2401	TGCCACTACGAGCTCTACTTCCCAATCACCAGAGGGTAATGCGAGGGATGGATAGCATCTCTCTTATTCTGCTG	
	C H Y E L Y F A N H Q R V M R G M D S I L L I L L	736
2476	GAACCTCTGCCGTTATATCTCATTCCTCCAAAGTACTACCAACTCAAGGCCATGATGTCAGACGCACATATCTA	
	E P L P L Y L I P S K Y Y Q L K A M M S R R T Y L	761
2551	GAGTGGCCACAGGAAGGCCAAGCAGAAGCTGTTTTGGGCAAAATCTCAGAGCAGCTCTGCAGGCCAACCTTCCC	
	E W P Q E G A K Q K L F W A N L R A A G L C A N L P	786
	ATTCCATCAGAGAGAGAAGAAGAGTGAAGAGAGGAGTTTACAAATTTATTTTCATAGAGATTGTCTTAGCATATA	
	I P S E R E E E *	794
2701	TTTATTTAGGCGAGTGGATCTCAACAAGGGTTCCAGGATCCACTAGGGGGCTCAACAAAACAAGATAAAACTAA	
2776	TCTAAAATCAAGATATTTCTAATTTTAATTCACCACCTCAACCTGGATACAAGAGGTAGCCTAATTTTAAAGTG	
2851	TGACTTCTAGTCTGGAATATTGAGGGATTAGCAAACTAAATAATATGTAATGGAATTCTTAAACATAAAAA	
2926	AAAAAAAAAAAAAAAAAAAA	

그림 1. 잉어 *tlr1* 유전자의 핵산과 아미노산배열

어두운 부분은 보존된 TIR구조배열, 한 밑선을 그은것은 로이신평부반복배열(LRR), 두 밑선을 그은것은 로이신평부반복배열의 C-말단배열(LRR CT), 점선을 그은것은 막투과단백질배열, 사체에 밑선을 그은것은 가능한 폴리아데닐산신호, 강조밑선을 그은것은 번역시작 코돈(ATG)과 종결코돈(TGA), ATTTA는 번역배열의 안정성을 특징짓는 배열이다.

BLAST 프로그램의 SwissProt자료기지를 리용하여 분석해보면 잉어*tlr1*유전자에 의하여 암호화되는 단백질의 분자량은 90.33kDa이고 등전점(pI)은 6.75이다.

그림 1에서 보면 이 유전자에는 번역시작코돈(ATG)과 종결코돈(TGA)이 있으며 3'-비번역배열에는 전사중지를 알리는 폴리아데닐신호(AATAAA)와 전사를 종결짓는 폴리A배열이 있다.

Expasy프로그램으로 분석하면 이 유전자는 신호펩티드배열(1-23aa)과 5개의 로이신풍부반복배열 (92~531aa), 1개의 C-말단로이신풍부반복배열(523-581aa), 1개의 세포막단백질투과배열(583-605aa)과 세포질안에 위치하는 1개의 TIR도메인(640-784aa)을 가지고있다.

TLR족 유전자는 세포질의 바깥에 병원체를 인식하는데 적응된 로이신풍부반복배열들을 여러개 가지고있으며 세포질의 안쪽에는 진화상 보존적이며 신호전달에 관여하는 TIR도메인을 가지고있다.[2]

2) 잉어*tlr1*유전자의 진화분석

진화분석에 의하면 잉어*tlr1*유전자의 배열은 물고기류(줄말고기)와 제일 가깝고 다음은 새류, 포유동물의 순서로 가깝다는것을 알수 있다.(그림 2)

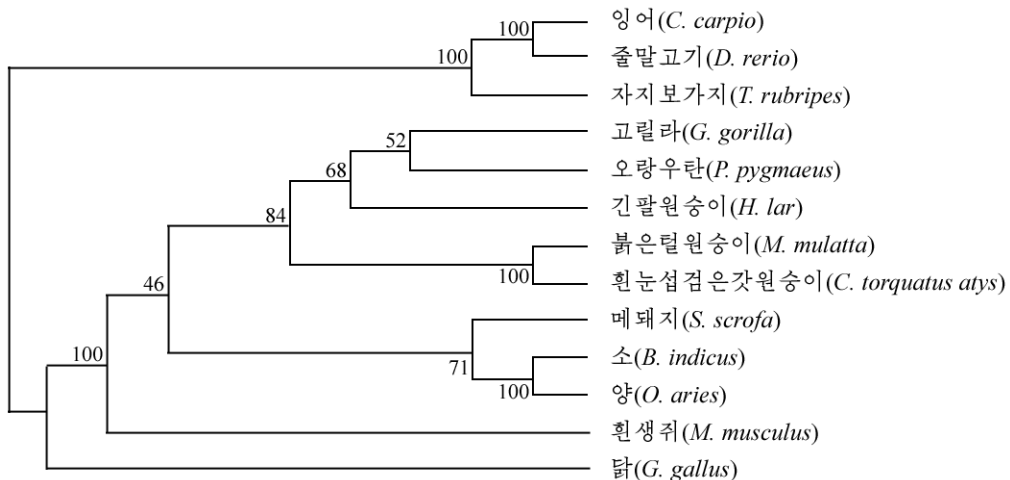


그림 2. *tlr1*유전자의 진화계통수

그림 2에서 보는바와 같이 이러한 진화관계는 진화규칙에도 완전히 부합되는데 이것은 유전자의 아미노산배열이 종의 진화를 과학적으로 반영하고있다는 증거로 된다.[3]

맺 는 말

1) 잉어*tlr1*유전자의 cDNA 전 배열의 크기는 2 948bp이다. 5'-비번역배열의 크기는 267bp이고 열린읽기틀(ORF)의 크기는 2 385bp이며 3'-비번역배열의 크기는 폴리A를 포함하여 296bp이다. 열린읽기틀은 794개의 아미노산을 암호화한다. 이 유전자는 5개의 로이신풍부반복배열과 1개의 막투과배열, 보존된 TIR도메인배열을 가지고있다.

2) 진화분석에 의하면 잉어*tlr1*유전자의 배열은 물고기류와 제일 가깝고 다음은 조류, 포유동물순서로 가깝다.

참 고 문 헌

- [1] J. R. Sambrook; Molecular Cloning, Cold Spring Harbor Laboratory, 114~140, 2001.
[2] S. Akira et al.; Cell, 124, 783, 2006.
[3] R. Medzhitov et al.; Nature, 388, 394, 2011.

주체104(2015)년 3월 5일 원고접수

**Analysis of the Sequence and Phylogenetic Relationships of
tlr1 Gene from Common Carp, *Cyprinus carpio***

Jang Song Hun, Paek Un Ju

The common carp *tlr1* has the full length cDNA consists of 2 948 nucleotides(nt). The 5'-untranslated region (UTR) is 267 nt. Open reading frame is 2 385 nt. The 3'-UTR is 296 nt, including a poly A tail. The ORF of *tlr1* gene encodes a protein composed 794 amino acids (aa). It has 5 LRR motifs; a C-terminal LRR motif, one transmembrane zone, and a TIR domains.

Analysis of evolutionary conservation showed that *tlr1* of common carp is much more close to the fishes than to the birds and then the mammals *tlr1*.

Key words: common carp, *tlr1* gene