

## 초어톨류사접수체유전자 5b(*tlr5b*)의 배열특성

손주성, 장성훈, 주창성

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《양어과학과 기술에 대한 연구사업을 강화하고 선진적인 물고기기르기기술을 적극 받아들여 우리 나라의 양어사업을 최신과학기술에 기초하여 발전시켜나가도록 하여야 합니다.》(《김정일선집》증보판 제20권 178~179페이지)

포유동물에서 TLR(Toll-like receptor)족 유전자는 병원체결합분자형태(PAMPs)를 인식하고 염증성세포인자들과 I형인터페론을 발현시켜 병원성세균에 대한 유기체의 면역을 조절한다.[1, 2] 우리는 초어를 연구대상으로 하여 물고기의 선천성면역에서 중요한 역할을 하는 *tlr5b*의 배열과 진화적특성을 밝혔다.

### 재료와 방법

실험에 리용한 프라이머 프라이머는 Primer premier 5.0프로그램으로 설계하였다.(표)

표 1. 초어*tlr5b* cDNA의 증폭에 리용한 프라이머

프라이머이름	배열(5'→3')	증폭위치
SMART II(TM)(정방향)	AAGCAGTGGTATCAACGCAGAGTACGCGGG	첫번째 사슬 cDNA의 증폭
CDS III(반대방향)	AAGCAGTGGTATCAACGCAGAGTAC(T)30V N	첫번째 사슬 cDNA의 증폭
gcT5b-F(정방향)	AAGGGTCTTGAGGCGAGTG	열린읽기를
gcT5b-R(반대방향)	CAAAGCAGCAGCGGAGTG	열린읽기를
T5bG1(반대방향)	CACTCGCCTCAAGACCCT	5'-말단
T5bNG1(반대방향)	TGTTGTACCCCAAGTAGAGC	5'-말단
T5bG2(정방향)	AGGGTCTTGAGGCGAGTGGT	3'-말단
T5bNG2(정방향)	GGTCTTGAGGCGAGTGGTAT	3'-말단

cDNA의 전 배열의 증폭 SMART cDNA의 합성방법[1]을 리용하여 초어의 두신으로부터 분리한 mRNA를 역전사시켜 주형cDNA를 얻는다. 다음 Vector NTI프로그램으로 다중 배열상동성을 검색하고 진화적으로 보존된 배열을 찾아내어 축퇴프라이머를 합성하였다. 이 프라이머를 리용하여 PCR법으로 초어*tlr1*의 보존된 배열을 증폭한다. 다음 증폭된 토막을 pMD18-T운반체(《Promega》)에 재조합하고 배열분석을 진행한다. 배열분석자료에 근거하여 5'-와 3'-말단을 증폭하기 위한 프라이머를 설계한다. SMART II와 T5bG1(T5bNG1), T5bG2(T5bNG2)와 CDS III프라이머로 목적유전자의 양쪽 말단을 증폭한다. 증폭된 5'- 및 3'-말단의 PCR산물들을 pMD18-T운반체에 재조합하고 그 배열을 분석하였다.

생명정보학적분석 핵산배열의 분석은 Vector NTI프로그램으로 진행하였다. 단백질의 기능령역에 대한 분석은 SMART와 TMHMM프로그램으로 분석하였다.

## 결과 및 논의

1) 초어 *tlr5b*의 핵산 및 아미노산배열에 대한 분석

SMART cDNA 합성방법을 리용하여 얻어낸 초어 *tlr5b*의 증폭산물들을 pMD18-T운반체에 클론화한 다음 배열분석을 진행하였다.

Vector NTI 프로그램으로 분석하면 3 324bp의 완성된 cDNA 배열(폴리 A 배열을 포함.)이 얻어졌다.(그림 1)

```

TAAACTTGGTTCTACATAAGTATTAACCTTAAATCGCACATCCTCTCAAAAATAGCCTCATG      60
TTTTTTTTTGTTTTTTTCCCCCAGTTCTCTCAGAATGGGATTACATTTATTCTGATCCTC      120
                                     M G F T F I L I L      9
TTTGGAATTGTGCATTAACACTGAAGTTGTGAAATGCACCTCTGTGTGTTTCAGTTAATGGC      180
 F G L C I N T E V V K C T S V C S V N G      29
TATGCAGCCTTCTGCATATCTAGAGGTCTTCATCAGGTGCCAGAGCTTCCACGTACATC      240
 Y A A F C I S R G L H Q V P E L P T Y I      49
AATTATGTGGATCTGAATTTTAAACAGCATTGCTGAACTCAACGAAACATCCTTTTCTCGT      300
 N Y V D L N F N S I A E L N E T S F S R      69
CTTGAAGGTCTACAAGTCCTTAAACTGGAGCAACAAACAACAGGACTTGTGATCAGAAAC      360
 L E G L Q V L K L E Q Q T T G L V I R N      89
AACACATTTAGAAGACTCTCCAATCTAATATTACTTAAAGTTAGACTACAACCACTTCTTG      420
 N T F R R L S N L I L L K L D Y N H F L      109
CGAATAGAGACAGGGGCATTTAATGGATTATTCAACCTTGAGATTCTCACTCTCACTCAG      480
 R I E T G A F N G L F N L E I L T L T Q      129
TGCAGTTTAGATGATACTATTTTGTCTGGTGACTTCCTCAAACCTCTGGTGTCTCTTGAG      540
 C S L D D T I L S G D F L K P L V S L E      149
ATGCTTGTCTTGCCTGAAAACAACATTAAGAAGATCCAGCCAGCATCGTTCTTTTAAAT      600
 M L V L R E N N I K R I Q P A S F F L N      169
ATGAGGAGATTCCATGTGCTCGATCTCTCTCGCAACAAAGTGAAGAGCATCTGTGAAGAA      660
 M R R F H V L D L S R N K V K S I C E E      189
GACCTCCTCAGCTTTCAGGGTAAACATTTACGCTTCTGAAGCTGTCCTCAGTGACACTG      720
 D L L S F Q G K H F T L L K L S S V T L      209
CAAGACATGAATGAGTACTGGTTAGGATGGGAAAAGTGTGGAAACCCATTTAAGAACATG      780
 Q D M N E Y W L G W E K C G N P F K N M      229
TCCGTAAGTGATTGGACTTATCTGGAAATGGCTTTAATGTTAAAATGGCAAAGCTTTTC      840
 S V S V L D L S G N G F N V K M A K L F      249
TTTGATGCAATCACTGGTACCAAAATCCAAAGTCTCATTTCTCAGTAACAGTTACAGCATG      900
 F D A I T G T K I Q S L I L S N S Y S M      269
GGCAGTTCTTTTGGTCATAACAATTTCAAAGATCCAGACAAATTTACTTTGAAGGGTCTT      960
 G S S F G H N N F K D P D K F T L K G L      289
GGGGATAGTGGTATTAAGATTTTCGATTTGTCCAAATCACAAATTTTGTCTTTGTCAAAT      1020
 G D S G I K I F D L S K S Q I F A L S N      309
TCAGTATTTAGTCATTTTCAAGATCTAGAACAAATTACATTGGCAGAAAATCAGATCAAC      1080
 S V F S H F Q D L E Q I T L A E N Q I N      329
ATTATTGAAATTGATGCATTTTGGGGTATGACAAATTTACTAAAGCTAAACCTGTCCAAA      1140
 I I E I D A F W G M T N L L K L N L S K      349
AACTTCCTGGGTAATATTGATTCTAATACATTTCAGAATCTAAAGAAGCTCGAGGTGCTT      1200
 N F L G N I D S N T F Q N L K K L E V L      369

```

GATTTGTCTTATAACCATATAAGGGTGCTTGGCGATAAATCGTTTCAGGGACTCCCAAGT	1260
D L S Y N H I R V L G D K S F Q G L P S	389
TTACTCAACTTAAATTTAACAGGAAATGCTCTTGAGTCAGTTCATGAATTTGCAACCCTA	1320
L L N L N L T G N A L E S V H E F A T L	409
CCTAACCTGAAGATAATCTACTTGGGTGAGAACAGAAATTTTCATCTTTGTCTAGTTTACCC	1380
P N L K I I Y L G E N R I S S L S S L P	429
AACATTGCTAAAAATCTCACAACCCTTGACCTGGAAATGAACAAATTACAGGCCTTGTC	1440
N I A K N L T T L D L E M N K L Q A L S	449
GATCTCTACACAATCTACGGGAATTTCTCAAATAGAAAAATCTTTCTTCAAGGTAAC	1500
D L Y T I L R E F P Q I E K I F L Q G N	469
AGTTTTTTCGAGTTGTTATAATCAAAGACAAATAGTGGCTTCAGACCAACTACAACCTTCTT	1560
S F S S C Y N Q R Q I V A S D Q L Q L L	489
CATCTCGGACGTTTCATCTATGCAGCTGATCTGGTCAGAAGGAAATGTTTAAATGTGTTT	1620
H L G R S S M Q L I W S E G K C L N V F	509
AACAATCTTCACCAGTTAGAACAGCTTTCTCTGACTGCCAATGGGCTACAGTCTCTTCCC	1680
N N L H Q L E Q L S L T A N G L Q S L P	529
AAAGACATTTTTTAAAGACCTTACCTCTTTGTCTTTTTTGGATTTGTCCTTCAACTCTTTG	1740
K D I F K D L T S L F F L D L S F N S L	549
AAGTACCTTCCAAACGGTATATTCCCTGAAAGTCTTCAAATTCTTAATCTTGAATATAAT	1800
K Y L P N G I F P E S L Q I L N L E Y N	569
TCTATTTATTTCAGTAGATCCAAATCTCTTTAGCACCCCTCAGCTACCTCAGCCTGATAAAA	1860
S I Y S V D P N L F S T L S Y L S L I K	589
AACGATTTCCGTTGTGATTGCAACCTAAGGGATTTCAAACATGGCTAAACCAAACCAAT	1920
N D F R C D C N L R D F K T W L N Q T N	609
GTAATCATTTTCTCACTCCATTGAGGATGTGATATGTGCCAGTCCTGAGGATCAGTACATG	1980
V I I S H S I E D V I C A S P E D Q Y M	629
GTTCCGTTGTGAGATCCAGCATACAATGTGAGGATGAAGAGGACGAGAGAAATGTTGAA	2040
V P V V R S S I Q C E D E E D E R N V E	649
AAACTGAGGCTTGTGCTTTTTATTTTCTGTACCGCACTTATCACGTTACTCACTGCTAGC	2100
K L R L V L F I F C T A L I T L L T A S	669
GCCATCATTTATGTCCGTCGACGTGGCTACATCTTCAAGCTTTACAAAAAATCATTGGC	2160
A I I Y V R R R G Y I F K L Y K K L I G	689
ACACTTGTGGATGGAAGCGAGAGGAGCCTGATCCTGACCAATTCTTATATGACGTGTAT	2220
T L V D G K R E E P D P D Q F L Y D V Y	709
CTCTGCTTTAGTTCCAATGATATTAAGTGGGTAGAAAGAGCACTGCTGAACAGGCTAGAC	2280
L C F S S N D I K W V E R A L L N R L D	729
TCTCAGTTCTCAGAGCAGAACACACTCCGCTGCTGCTTTGAGGAGCGAGACTTCATACCC	2340
S Q F S E Q N T L R C C F E E R D F I P	749
GGGGAGGACCATCTTACCAACATGCGAAATGCTATCCAGAATAGTCATAAAACCTTTGT	2400
G E D H L T N M R N A I Q N S H K T L C	769
GTGGTGTCTGAACATTTCTTGAAGGATGGCTGGTTACTAGAGGCCTTCAATTCTGGCACAA	2460
V V S E H F L K D G W L L E A F I L A Q	789
AGAAGGATGCAAGTGGAGCTTGAGGACATTCTGGTGGTGCTGGTTGTAGGGAACATACCG	2520
R R M Q V E L E D I L V V L V V G N I P	809
CAGTACAGGCTACTGAAGTTCAAACAAGTGAGATCCTACATTGAGAACAGAAGATACCTT	2580
Q Y R L L K F K Q V R S Y I E N R R Y L	829
CTGTGGCCCGATGACAGCCAGGACTTGGATTGGTTTTATGACCAACTTCTGCATAAAATA	2640
L W P D D S Q D L D W F Y D Q L L H K I	849
AGAAAAGATACCAAGGTTAAACAAACAAATCAACCAACCAACAAACAAAGCCTGAAGCA	2700
R K D T K V K Q T N Q P T K Q T K P E A	869

TTGAATGTCCATGCAAACACATCAGTATA <b>TAA</b> TTATGATATATCTGTTTATTATATATTTTG	2760
L N V H A N T S V *	879
TAAAAGTTTACCTAATTACCTTAGATTTTTCAGATTGTTACAATGCATAATGCAGATCTTG	2820
GTTAGGTATAATTATGCATTATCTCATGTTTCTATGTTTGTAATAATGTTTGTCTGTAA	2880
ACCTGCTGATAATTATAGTAAATATCTGAAATGCTAATTCTAGCAGAATGCGTGTCCTGC	2940
TGCCAGTGGATAAAAGACTCTGCTAAATTTGCAAAATCTTAACATCTTCAACCACCTGTGT	3000
GATTTTGACAATTGCAAACTCTT <b>AATA</b> AAATGTGAATTACAAATGAAGAAATTAAGCAAA	3060
TTAC <b>ATTTA</b> ACTATGA <b>ATTTA</b> CAACCTAAAAATGTTAACTGGATAAAAAAGACACTCAGTG	3120
TACTCTGACCTTATCTTATCTGAATACAGGCAAAATGGCAAAGCCACTAGTCA	3180
AAACACTGACACAATCTGGCCTTATGCTTTTGGGTAATCACAGTTTGTACATTGTTAATG	3240
AAATCTGTAAAGAAACCTCAGTATTGACAAATCTGTGCATGGCCTGTGTACTTTGAACAG	3300
TTTGGAATGTAGCAATAGGGCTGAAAAAAAAA	3424

그림 1. 초어*tlr5b*의 핵산과 아미노산배열

테두리를 친 부분은 보존된 TIR구조배열, 1개의 밑선을 그은것은 로이신평부반복모티프배열(LRR), 2개의 밑선을 그은것은 로이신평부반복모티프(LRR)의 C-말단배열, 회색으로 표시한 부분은 세포막투과배열, 사선강조체로 표시한것은 가능한 폴리아데닐산신호배열, 강조체에 굵은 밑선으로 표시한것은 전사배열의 불안정성을 보여주는 모티프배열, 번역개시코돈(ATG)과 중지코돈(TAA)은 강조체로 표시

초어*tlr5b*에서 5'— 및 3'—비번역배열의 길이는 각각 93, 594bp이고 열린읽기틀의 길이는 2 637bp이며 3'—비번역배열에 배열의 불안정성을 보여주는 2개의 모티프(ATTTA)들이 있다. 이 배열에 의하여 암호화되는 단백질의 크기는 878aa이다.

그림 1에서 보면 이 유전자에는 전사시작코돈(ATG)과 중지코돈(TAA)이 있으며 3'—비번역배열에는 전사중지를 알리는 폴리아데닐신호배열(AATAAA)이 있다.

## 2) 초어TLR5b접수체단백질의 기능영역분석

SMART와 TMHMM프로그램으로 분석하면 초어TLR5b접수체유전자는 1개의 신호펩티드배열(1~21aa)과 11개의 로이신평부반복배열(95~581aa), 1개의 C-말단로이신평부반복배열(590~640aa), 1개의 세포막투과배열(653~675aa)과 세포질안에 위치하는 1개의 TIR도메인(705~856aa)을 가지고있다.(표 2, 그림 2)

표 2. 초어TLR5b접수체에서 아미노산배열의 기능영역배열

배열이름	부터(aa)	까지(aa)	E-값
신호펩티드(signal peptide)	1	21	N/A
로이신평부반복모티프(LRR)	95	118	0.867
로이신평부반복모티프(LRR)	119	144	309
로이신평부반복모티프(LRR)	145	168	0.996
로이신평부반복모티프(LRR)	315	338	4.57
로이신평부반복모티프(LRR)	339	362	1.07
로이신평부반복모티프(LRR_TYP)	363	386	0.001 95
로이신평부반복모티프(LRR)	387	408	14.5
로이신평부반복모티프(LRR)	409	434	15.7
로이신평부반복모티프(LRR_TYP)	512	535	0.016 7
로이신평부반복모티프(LRR_TYP)	536	559	0.041 1
로이신평부반복모티프(LRR)	560	581	29.2
C-말단로이신평부반복모티프(LRRCT)	590	640	$3.29 \cdot 10^{-11}$
세포막투과영역	653	675	N/A
톨인터로이킨접수체도메인(TIR)	705	856	$9.38 \cdot 10^{-19}$



그림 2. 초어TLR5b접수체의 기능영역

포유동물의 TLR5접수체는 세포질의 바깥에 세균의 편모를 특이적으로 인식하는데 적응된 로이신풍부반복배열들을 여러개 가지고있고 세포질의 안쪽에는 진화상 보존적이며 신호전달에 관여하는 TIR도메인을 가지고있다.[2]

### 맺 는 말

1) 초어*tlr5b*의 cDNA전배열의 크기는 3 324bp이다. 5'-비번역배열의 크기는 93bp이고 열린읽기틀(ORF)의 크기는 2 637bp이며 3'-비번역배열의 크기는 594bp이다. 열린읽기틀은 878개의 아미노산을 암호화한다.

2) 초어TLR5b접수체단백질은 11개의 로이신풍부반복배열과 1개의 막투과배열, 보존된 TIR도메인배열을 가지고있다.

### 참 고 문 헌

[1] S. Giovanna et al.; Genomics, **64**, 230, 2000.

[2] S. I. Yoon et al.; Science, **335**, 859, 2012.

주체105(2016)년 1월 5일 원고접수

## Sequence Characteristics of the *tlr5b* Gene from Grass Carp, *Ctenopharyngodon idella*

Son Ju Song, Jang Song Hun and Ju Chang Song

The grass carp *tlr5b* gene has the full length cDNA consisting of 3 324 nucleotides (nt). The 5'-untranslated region (UTR) is 93nt and open reading frame is 2 637nt. The 3'-UTR is 594nt. The ORF of grass carp *tlr5b* gene encodes a protein composed of 878 amino acids (aa). It has 11 LRR motifs, two C-terminal of LRR motif, one transmembrane zone and a TIR domains.

Key words: grass carp, *Ctenopharyngodon idella*, *tlr5b*, gene sequence, LRR motif