

## 초어톨류사접수체유전자 3(*tlr3*)의 배열특성

로진성, 장성훈

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《기초과학부문들을 발전시켜야 나라의 과학기술수준을 빨리 높일수 있고 인민경제 여러 분야에서 나서는 과학기술적문제들을 원만히 풀수 있으며 과학기술을 주체성있게 발전시켜나갈수 있습니다.》(《김정일선집》 증보판 제10권 485페이지)

유전자의 분자구조적특성을 밝히는것은 해당 유전자의 기능을 깊이있게 해명하고 실 천에 도입하는데 필요한 기초자료를 마련하는데서 큰 의의를 가진다.

우리는 초어의 선천성면역조절에서 중요한 역할을 하는 *tlr3*의 cDNA배열에 대한 연 구를 하였다.

### 재료와 방법

#### 1) 초어*tlr3* cDNA 전 배열의 증폭

SMART cDNA의 합성방법[1]을 리용하여 초어의 두신으로부터 mRNA를 분리한 다음 설명서에 따라 cDNA 전 배열을 얻었다.

표 1. 초어 *tlr3* 증폭에 리용된 프라이머

프라이머이름	배열(5'- 또는 3'-)	증폭길이(nt)와 용도
SCTF1(상류프라이머)	CTCAGYAAYAAYAAYATYGCMAACAT	668, 유전자클론화 (바깥)
SCTR4(하류프라이머)	TTCCAGWAGAAYYGRATYCTCCA	
SCTF2(상류프라이머)	TKCAGCACAAAYAAYTTDGC	360, 유전자클론화 (내부)
SCTR3(하류프라이머)	AYAGRATGCTCTCRCASGTGCA	
SCTF20(상류프라이머)	CAAAACTGTCTGTTCTGGATCTG	3'-말단
SCTF6(상류프라이머)	TGATGCTTTGCGTGGCTTCTCTGA	
SCTR22A(하류프라이머)	TCAGAGAAGCCACGCAAAGCATCA	5'-말단
SCTR23A(하류프라이머)	CAGATCCAGAACAGACAGTTTTG	
SCFA(상류프라이머)	GAAAAAGCTTGGAAGAAAAGAAAA	3 467, 내부배열
SCRA(하류프라이머)	AGGCTGAAACAATGAAACAATAAA	
UPM 긴배열	CTAATACGACTCACTATAGGGCAAGCAGTGGT	말단증폭
UPM 짧은배열	ATCAACGCAGAGT	
NUP	CTAATACGACTCACTATAGGGC	5'-말단
5'-RACE 올리고dG	AAGCAGTGGTATCAACGCAGAGT	
3'-RACE 3'-CDS	AAGCAGTGGTATCAACGCAGAGTAC (T) 30VN	3'-말단

R=A/G, S=C/G, Y=C/T, D=A/G/T, K=G/T, M=A/C, W=A/T, V=A/G/C, N=A/G/C/T

#### 2) 배열분석결과의 처리

DNA배열은 Vector NTI와 DNA star, Chromas프로그램으로 분석하였다.

단백질배열은 Expasy로, 단백질의 기능영역은 SMART, TMHMM프로그램으로 분석하였다.

## 결과 및 논의

SMART cDNA 합성 방법을 리용하여 5'-비번역배열과 열린읽기틀, 3'-비번역배열, 폴리A를 포함한 3 681bp의 초어 *tlr3* cDNA배열을 얻어냈다.(결과는 생략) 완성된 초어 *tlr3* cDNA배열에서 5'-비번역배열의 크기는 148bp이고 열린읽기틀(ORF)의 크기는 2 715bp이며 3'-비번역배열의 크기는 758bp이다.(그림) 초어 *tlr3*에서 계놈배열과 cDNA배열을 비교해보면 이 유전자의 구조는 1개의 엑손으로 되어있다는 것을 알 수 있다.

```

1  GTTTGTTCCTTCTTAAGAAAGAGGAATTGTTCCCTGTTAGATATTTTCCCCCTGCCAA
61  GACTATAGGGTTTACAAGAAGACTGACTGGATTGCTTCTATTGTAACTACTAATCCATT
121 GTGGGATATTGAGCTGTAAGGTGGGAAATAGATGGTCAGAGTTCATTTTAACTTGGGGG
1  M E L M K L I L L P L
181 AAAAAGCTTGGAAGAAAAGAAAATATATATGGAAGCTGATGAACTCATACTGTTGCCCTT
12  F Y T C F H A H C A G S A Y P H R S T C
241 GTTCTACACTTGCTTCCATGCCACTGTGCAGGCTCTGCATATCCACACAGGTCAACATG
32  T I E N A K A D C S H M N L D V V P T N
301 TACGATTGAAAATGCCAAAGCAGACTGTAGTCATATGAATCTGGATGTAGTTCCAACGAA
52  L P K N I T T L D V S H N R L K N L S S
361 TTTACCCAAAATATCACCACATTGGATGTGTCTCACATAGACTAAAAATCTGTCTTC
72  L L L Y S N L V N V D A S Y N S L T A I
421 CTGTCTTTGTACTCGAATCTGGTGAATGTAGATGCCAGCTACAACCTCTTAACTGCCAT
92  E K D L C L S L P H L Q I L N V Q H N E
481 AGAAAAGGATCTATGTCTTTCTGCCACACCTGCAAATCTTAATGTGCAACACAATGA
112 V Y L M N E K V L K N C F H L M R L D L
541 AGTGTATTTGATGAATGAGAAAGTCTGAAAACTGTTTTCATTTGATGCGACTTGACCT
132 S G N R L K L Q G E P F S V L K N L T W
601 GTCTGGCAATAGGCTGAAGCTACAAGGGAGCCTTTCTCTGTTCTAAAGAATTTGACATG
152 L D V S R N K L K S A K L G T Q P Q L P
661 GTTGGATGTATCTCGAAACAACTGAAATCTGCTAAACTGGGCACACAACCTCAGCTGCC
172 N L V T L I L S G N E I S A L Q K N D F
721 AAACCTGGTGACCCTTATCTTTCTGGAATGAAATTTCTGCACTGCAAAAGAATGACTT
192 S F L S N S A F R V L I L S G S L S L K
781 CTCATTCCTAAGTAATTCCTCTGCATTTCTGGGTTCTGATACTCTCGTCTCTGTCTCTTAA
212 K V E N G C F Q A I S R L S D L V L D W
841 GAAGGTCGAGAAATGGCTGTTTCCAGGCTATTTCTAGACTATCTGACTTGGTCTCTGGATTG
232 S R L T Q F T N L C E E L A G T A L
901 GAGCAGGCTCAGCACTCAGTTTACCACTAATCTTTGTGAAGAACTTGCTGGCACAGCCTT
252 R N L S L K N T F Q V T L S D T T F K G
961 GCGAAACCTTTCCCTTAAGAACACTTTCCAGGTGACACTCTCAGACACGACTTTTAAAGG
272 L D K T N I T V L D L S Y N T M S K I A
1021 TCTAGACAAGACTAACATCAGTGCTTGACCTCAGCTACAACCAATGTCTAAGATTGC
292 D G A A F Q W F P R L E F L S L E H N S L
1081 TGATGGTGCCTTTTCAGTGGTTTCCCCGACTGGAATTTTTATCCCTGGAGCATAACTCCCT
312 R H L T K D T F S G L G N L R Q L N L Q
1141 TAGACACCTAAGTAAAGGACACCTTTCAGTGGACTGGGAAACCTGAGGCAGCTTAACCTGCA
332 K A A L I K S H T S S L P I I E D F S F H
1201 GAAAGCACTGATTAAGAGTCATACATCATCTTGCCAAATTATTGAAGACTTCTCATTTCA
352 H L V Q L E H L C M A N T A F R E I T E
1261 CCATTTAGTCCAATTGGAGCATCTGTGTATGGCAAATACTGCATTCCGAGAGATAACAGA
372 H T F S G L L H L K T L D L S W S S T G
1321 GCACACCTTCTCTGACTTCTGCACCTGAAGACACTGGATTAAAGTTGGAGCAGCAGG
392 L K T V T N K T F A S L Q E S P L L E T
1381 GTTGAAAACAGTCACAAACAAAACCTTTGCTTCTCTGCAAGAATCACCACTCCTTGAGAC
412 L N L T A M G I N K L G P G A F S S L G
1441 TCTTAATCTTACAGCCATGGGTATAAACAAGTTGGGGCCTGGTGCCTTCTCAAGTTGGG
432 N L T T L L L G R N F I N Q Q L R G D E
1501 CAACCTCAGTACACTCTACTTGGCCGAACTTCATTAATCAGCAGTTGAGAGGGGATGA
452 F E G L T S I K E I D M S I N Q Q S I S

```

1561 GTTTGAGGGCCTAACTAGCATTAAAGAGATTGACATGTCTATAAACCCAGCAAAGTATTTTC  
 472 L T N T S F I H V P T L R T L K L G R A  
 1621 CCTTACAAATACGTTCATCCACGTCCCTACGTTGAGGACTCTAAAGCTAGGTCGTGC  
 492 L K G T C L D I E P S P F R P L V N I T V  
 1681 CCTAAAAGGGACCCTAGATATTGAACCATCTCCATTGAGGCCACTAGTCAACATCACAGT  
 512 L D L S N N N I A N I N D G M L R G L Y  
 1741 TCTAGATCTCAGTAATAACAATATTGCAAAACATAAATGATGGCATGCTGAGAGGGCTGTA  
 532 H L K V L K M Q H N N L A R L W K T A N  
 1801 TCATCTGAAAGTGCTGAAAATGCAGCACAACTTGGCTAGGTTGTGGAAGACGGCCAA  
 552 P G G P V F F L K D A T K L S V L D L D  
 1861 CCCTGGTGGTCCAGTGTTTTCTCCTCAAGGATGCCACAAAACCTGTCTGTTCTGGATCTGGA  
 572 Y N G L D E I P L D A L R G F S E L H E  
 1921 TTACAATGGCCTAGATGAGATTCCACTTGATGCTTTGCGTGGCTTCTCTGAGTTACATGA  
 592 L S L R G N L L D Q L H A S V F F D D L Q  
 1981 GCTAAGTCTCCGTGAATCTCTTGATCAACTGCATGCCTCTGTTTGTGATGACCTACA  
 612 S L K Y L H L Q K N L I T S V Q R A T F  
 2041 GTCCTTAAAGTATTTGCATCTTCAAAAGAACCTTATAACGTCCGTCCAACGTGCCACGTT  
 632 G V P L S N L T E L Y M D R N P F D C T  
 2101 TGGTGTGCCCTTGTCCAACCTGACAGAACTTTACATGGACCGCAACCCCTTTGACTGCAC  
 652 C E S I L W F S E W L N S T N S V P G  
 2161 CTGCGAGAGCATCCTGTGGTTTCCGAGTGCGTTAACTCAACCAATGCAAGCGTTCCTGG  
 672 F P Q S Y I C N T P N A Y F N R S V M Y  
 2221 ATTTCTCAAAGTTACATCTGCAATACCCCAAATGCCTACTTTAAACCGTCCGTTCATGTA  
 692 F D P L S C K D M T P F K A L Y I M T S  
 2281 CTTTGACCCATTGTCTGCAAGGATATGACACCTTTTAAAGGCACTGTACATCATGACTAG  
 712 T A V L M L L F T A F L V H F Q G W R I  
 2341 CACAGCAGTTTTAATGTTGTTATTACAGCTTTTCTGGTGCCTTCCAGGGATGGAGAAT  
 732 Q F F W N I I V N R M L G S P K D E S V  
 2401 CCAGTTCTTCTGGAACATAATAGTAAACCGTATGCTGGGGTCGCGCAAGGATGAAAGCGT  
 752 T E G Y D A Y I I H T A K D R P W  
 2461 TACTGAAGGTAGATATGTGTATGCTTATATCATTACACTGCTAAGGACAGACCATG  
 772 V E R S L L P L E D E N F N F F L E D R  
 2521 GGTGGAACGAAGCTTGCTCCCCCTAGAGGATGAAAATTTTAAATTTTTTCTTGAAGACAG  
 792 D A I P G F S Q L N T I I E N M G Q S R  
 2581 AGATGCAATACCTGGCTTTTCTCAACTTAACACCATTATTGAAAACATGGGACCAATCCAG  
 812 K I I F V I T E M L L K D P W C R Q F K  
 2641 GAAGATCATTTTTGTTATCACAGAAATGCTTTTAAAGGATCCATGGTGTAGGCAATTCAA  
 832 A H H A L H Q V M E N N R D S L I L I F  
 2701 AGCATCATGCACTTACCAGGTAATGGAGAACAATCGTGACTCCCTGATCTTGATTTT  
 852 L Q D V T D Y N L N R S L Y L R R G M L  
 2761 CTTGCAAGGTAACTGATTACAATTTGAACCGCTCTCTGTATCTTCGTCGTGGCATGTT  
 872 K P R C V L Y W P L H R E R I P A F H Q  
 2821 GAAACCTCGCTGTGTTCTCTACTGGCCTTTACACAGGGAACGTATCCCAGCTTTTTCATCA  
 892 K L R S A L A S S N K V N \*  
 2881 GAACTCCGCTCAGCATTAGCCTCTTCCAACAAGGTTAATTAGAAATGATTTTTCCATAA  
 2941 TGTTTATTTTGTGCCAGGTCTCAATATTTGTACTACAGTCTTACTTGGATTCTCTGTT  
 3001 CCATCACTTTAACATTTCTTAACACATGTACAGTTTAAAAGAATTTTCAGGTCAGATAAG  
 3061 GTGGAGCACTGGTTCTTTTCAAATCCTGCAGACCATGTACCTACCAGAAGTTGCATCCA  
 3121 ATGGCTTTGCAAGGAAGCTACCAAATGTGACCTGACAAATTTATTTTAAACAGATAAATTT  
 3181 ACTGTGCTCACATGAGAAGTTTCATTACCTTTTTTTTTTTTATTTTTTTTTTATCGTCTT  
 3241 TAGTAACTATAATACATAATCCATTTTCAAACTCCAACTAGACACTTGGGCAGATGCAA  
 3301 TAAGGCCCAATCTGTCTTAGGTGCTATTTTCTTATGCCCCAAAGAGATCCTTTTACACAAA  
 3361 CTAGCATTTGTAACCATAACGATGCAATTTAATCTGGCCATTTGCTGGCCACTGCCATCA  
 3421 GAGCCACAATATAAGGTAAATAGGACCTATGTAATTTAGCAATCACTTTCAAGGATCAAG  
 3481 GAACTAAAAATATTAATTTTATCATTAAATTAATTTTCCCTCATGAGTTTATTAATAG  
 3541 TTTTAAATCCACAAATACAGCAGGGGGATACGTTTTTGTGAAGTTGAGTACTTTTGG  
 3601 CATGTTGGCAGGTACTAATTATTTTATTGTTTTCATTGTTTCAGCCTGAGAAGGGGTAATT  
 3661 AGTAAAAA

그림. 초어*tlr3*의 핵산과 아미노산배열

물결표로 표시한 부분은 보존된 TIR구조배열, 밑선을 그은것은 로이신풍부반복배열(LRR), 2개의 밑선을 그은것은 막투과배열, 회색으로 표식한 부분은 신호펩티드배열, 사선강조체에 밑선을 그은것은 가능한 폴리 아데닐 산신호배열(polyA signal), 사선강조체는 ATTTA배열, 번역개시코돈은 네모칸에, 종결코돈은 \*로 표시함

이 유전자에 의하여 암호화되는 단백질은 904개의 아미노산으로 되어있다. 3'-비번역배열에는 폴리아데닐산신호배열(AATAAA)이 있고 mRNA배열의 불안정성을 보여주는 4개의 ATTTA배열들이 있다. 선행연구[2]에 의하면 이 배열은 mRNA가 빠른 속도로 분해되게 하는 기능을 수행한다. Blastp분석에 의하면 번역된 배열이 이미 밝혀진 TLR3단백질들과 상동성이 매우 높다.

초어 *tlr3*에 의하여 암호화되는 904개 아미노산으로 이루어진 TLR3단백질의 분자량은 102 765Da이며 등전점은 8.35이다. Blastp분석에 의하면 이 유전자는 유럽붕어의 TLR3 유전자와 상동성이 제일 높다.(E값 0.0) 그리하여 이 유전자의 이름을 초어에 있는 *tlr3*의 상동유전자라는 의미에서 *CiTLR3*으로 명명하였다. *CiTLR3*은 1개의 신호펩티드, 14개의 로이신풍부반복배열(LRR), 1개의 막투과배열, 1개의 톨인터로이킨접수체도메인을 가지고있다.

TLR족 유전자는 세포질의 바깥에 병원체를 인식하는데 적응된 로이신풍부반복배열들을 여러개 가지고있으며 세포질의 안쪽에는 진화상 보존되고 신호전달에 관여하는 TIR도메인을 가지고있다.[3, 4]

초어 *tlr3*의 구조에 대한 연구결과들은 물고기류에서도 포유동물에서와 마찬가지로 *tlr3*이 비루스를 인식하고 신호통로를 통하여 선천성면역반응을 일으키는 기능을 수행할 수 있다는것을 보여준다.

## 맺 는 말

1) 초어 *tlr3* cDNA의 전 배열크기는 3 681bp이며 5'-비번역배열과 열린읽기틀, 3'-비번역배열의 크기는 각각 148, 2 715, 758bp이다.

2) 이 유전자는 904개의 아미노산을 암호화하며 분자량은 102 765Da이고 등전점은 8.35이다.

## 참 고 문 헌

- [1] J. R. Sambrook; Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, 11~25, 2001.
- [2] A. E. Kel et al.; Nucleic Acids Res., 31, 3576, 2003.
- [3] J. Lee; Curr. Opin Gastroenterol., 23, 27, 2007.
- [4] X. Jiang et al.; Immunity, 36, 959, 2012.

주제105(2016)년 11월 5일 원고접수

## Sequence Characteristics of the *tlr3* Gene from Grass Carp *Ctenopharyngodon idella*

Ro Jin Song, Jang Song Hun

The full length of the grass carp *tlr3* is 3 681 nucleotides(nt). The 5'-untranslated region(UTR) is 148nt and open reading frame is 2 715nt. The 3'-UTR is 758nt including a poly A tail. The 904 amino acid polypeptide encoded by this gene has a calculated molecular weight of 102 765 Da and isoelectric point of 8.35.

Key words: grass carp, *Ctenopharyngodon idella*, *tlr3*, gene sequence