

초어의 핵인자유전자(*Cinfbk*) cDNA의 배열특성

안준혁, 장성훈

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

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

유전자의 분자구조특성을 밝히는것은 해당 유전자의 기능을 깊이있게 해명하고 실천에 도입하는데 필요한 기초자료를 마련하는데서 큰 의의를 가진다.

우리는 초어의 선천성면역조절에서 중요한 역할을 하는 핵인자유전자(*nfkb*)의 cDNA 배열에 대한 연구를 하였다.

재료와 방법

1) 초어*nfkb* cDNA 전 배열의 증폭

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

초어*nfkb*유전자증폭에 리용된 프라이머들을 표 1에 주었다.

표 1. 초어*nfkb*유전자증폭에 리용된 프라이머

프라이머이름	배열(5'→3')	크기/bp	증폭단편크기/bp	$T_m/^{\circ}\text{C}$
<i>nfkb</i> -IF	GTTTCCAACCCTATCTATGA	20	195	62
<i>nfkb</i> -IR	GAAGTCTCCAAATGCCTC	18		
02F23	GCCCAAGTTAGTTGAGGTGCTG	22		
02F24	GCACTGCTCGGGCTTTGTTACAG	23	1 215	58.2
02nkF	CCTGTTGTTTCCAACCCTATCT	22		
02nkR	CTCACCCCTCATCATCTTTCATCT	22		
02R23	TCCTGTGACGTGCTGTATTTCC	22	1 491	55.8
02R24	TGTGCTGGGGAAATGACAGTTGC	23		
nf-5F	TGGTGGACCTTCCATATCTCC	21		
nf-5R	CAACAGGCTTCAAGGCTCT	19	606	56.4
nf-6F	AGAGCCTTGAAGCCTGTTG	19		
nf-6R	TGTTTGTGGACGTCAGTGGG	20		
nf-7F	ATCAGTGCTGGGAGGAGAAG	20	236	56.3
nf-7R	TGATGAACCTTCTGCTGTGGG	20		
nf-8F	GGTGCAACAGCTTATTCAGG	20		
nf-8R	CAGCATGGAGCAGAGAGGTG	20	547	57.5
nf-9F	ACGTGAATGTGTGTACCTACGG	22		
nf-9R	TCCTGGAACAGATGAGTGAGTGT	23		
nf-IF	GAGACTTCTGACAGCCAGCCAAC	23	520	57.8
nf-OF	ATCAAGGTTTGGACAGCGAGG	21		

2) 배렬분석결과와 처리

DNA배렬은 Vector NTI와 DNA star, Chromas프로그램으로 분석하였다.

단백질배렬은 Expasy로, 단백질의 도메인영역은 SMART, TMHMM프로그램으로 분석하였다.

결과 및 논의

SMART cDNA합성방법을 리용하여 5'—비번역배렬과 열린읽기틀, 3'—비번역배렬, 폴리A를 포함한 3 351bp의 초어 $nfkb$ 유전자cDNA배렬을 얻어냈다.(결과는 생략) 완성된 초어 $nfkb$ cDNA배렬에서 5'—비번역배렬의 크기는 3bp이고 열린읽기틀(ORF)의 크기는 2 739bp이며 3'—비번역배렬의 크기는 609bp이다.(그림 1)

```

tcaATGGCTGGAGCACTAAGGATGGATGATGCGCAATATCCAATGAAGTTTGACAATGAG 60
  M A G A L R M D D A Q Y P M K F D N E 19
TCTTTACAGTTTATGGTGGACCTTCCATATCTCCATGAACCTTTTGAAGTGAAGCATGAA 120
  S L Q F M V D L P Y L H E P F E V K H E 39
CCCTTTGTGACAGAAACAGTTCATGGGCCTTACCTACAGATAATTGAAGAGCCAAAGCAG 180
  P F V T E T V H G P Y L Q I I E E P K Q 59
AGAGGATTCAGATTTTCGTTATGAATGTGAGGGCCCATCCACGGTGGTCTGCCAGGGGCC 240
  R G F R F R Y E C E G P S H G G L P G A 79
TCCAGTGAGAGGAACAGAAGGACGTATCCACAGTCAAGGTTTCAAACATATGTGGGTCAT 300
  S S E R N R R T Y P T V K V S N Y V G H 99
GCTCGCGTGGAGGTTTCAAGTGTGACACACACAGACCCCCCTCGTGTTTCATGCGCACAGC 360
  A R V E V Q L V T H T D P P R V H A H S 119
CTTGTGGGTCGTCAGTGTAAACGACAACGGGATATGCAGCATTTGATGTTGGGCCTAATGAT 420
  L V G R Q C N D N G I C S I D V G P N D 139
CTCACAGCACAATTCAGTAATTTAGGAATCCTCCATGTTACTAAGAGAGGTGTGGCTGAA 480
  L T A Q F S N L G I L H V T K R G V A E 159
GTATTAACAAAAAGGCTGAGAGAAGAGAAGAGGAGGACAAAGGAGCCTGGATATAAGTTT 540
  V L T K R L R E E K R R T K E P G Y K F 179
AGTGATGCAGAGGAACAGGCTATTATGCAAGAAGCCAAAGAATTGGGCAAAATATGGAC 600
  S D A E E Q A I M Q E A K E L G K N M D 199
CTAAACATTGTAAGGTTAAAGTTCAGTGTACCTTACCTTACAGGACAGCAATGGGAGTTTCA 660
  L N I V R L K F T A Y L Q D S N G S F T 219
AGAGCCTTGAAGCCTGTTGTTTCCAACCTATCTATGATAGCAAATCACCAATGCTTCA 720
  R A L K P V V S N P I Y D S K S P N A S 239
AATCTGAAGATCTCCCGAATGGATAAAACCTCTGGATCAGTGCTGGGAGGAGAAGAGGTC 780
  N L K I S R M D K T S G S V L G G E E V 259
TTTCTGCTCTGTGATAAAGTTCAAAAAGATGATATTGACATCCGTTTTTTATGAGGAGGAA 840
  F L L C D K V Q K D D I D I R F Y E E E 279
GATGATTCGCGATGGGAGGCATTTGGAGACTTCTCCCCACTGACGTCCACAAACAGTAT 900
  D D S G W E A F G D F S P T D V H K Q Y 299
GCGATCGTGTTTAAGACACCACCGTACCGCTGCACAGATATCACACGCCCTGTACAGTG 960
  A I V F K T P P Y R C T D I T R P V T V 319
TTCCTGCAGCTCAAGAGGAAGAAGGGAGGCGACTGCAGCGAACCCAAAGCAGTTACATAT 1020
  F L Q L K R K K G G D C S E P K Q F T Y 339
ATTCCACACAATCAAGACAAAGAAGTTTCAAGAGGAAGAGGATGAAGGCACTTCCTCAG 1080
  I P H N Q D K E E V Q R K R M K A L P Q 359

```

CACTACGATAACTGGCGGGGAGGACCTCAAGGAGGAGCAGGGGCCTTTGGAGGTCCTGGA 1140
H Y D N W R G G P Q G G A G A F G G P G 379
TCTGGAACAGGAGGAGGTGGAATGGGAGGAGTTTTCAATTTAATCATCCGATGGATGGT 1200
S G T G G G G M G G G F Q F N H P M D G 399
TCAGGGTTTTTTCATTGGTGGATGTGGTGGATTTGGTGGAGGGGCACAGATGTCAGGCTCC 1260
S G F F I G G C G G F G G G A Q M S G S 419
ACCCACAGGCTGATGGATCTCAGAATCAGCCAACAAACAAACAAAATCAGTCTAGCTCT 1320
T P Q A D G S Q N Q P T N K Q N Q S S S 439
CAAATGCAACAGCAGCTCATTGAGATTGCCACAGTTCTGCAGAATCGTGCCACTTTCGATG 1380
Q M Q Q Q L I Q I A T V L Q N R A T S M 459
GCAAAGGGCACTGCTCGGGCTTTGTTACAGTACTGCAGCACAGGGGATGTGCGCCCCCTG 1440
A K G T A R A L L Q Y C S T G D V R P L 479
CTGGTTCTGCAGAGACATCTGTGTGGAGTGCAAGATGAGAACGGAGACACGCCTCTGCAC 1500
L V L Q R H L C G V Q D E N G D T P L H 499
TTGGCCATTATCCATCAGCAACCTGCTGTGGTGAACAGCTTATTCAGGCCCTGAACAAC 1560
L A I I H Q Q P A V V Q Q L I Q A L N N 519
ACCCACAGCAGAAGTTCATCAACAACTCAACAACCTTAGTCAGTCTCCATTGCATCTG 1620
T P Q Q K F I N K L N N L S Q S P L H L 539
GCTGTGATCACTAAGCAGCCCAAGTTAGTTGAGGTGCTGATGAGAACGGGGCAGACCCG 1680
A V I T K Q P K L V E V L M R T G A D P 559
AGCCTGCAGGACCGAGAGGGCCGGACAGCGCTTCACTTGCCGCACACACCGGCGATGAG 1740
S L Q D R E G R T A L H L A A H T G D E 579
GCTGTACTACGAGTGGTCCTGGGCCTGCTTGGAGAGCGCTATATCCATTT**AATAAA**CTCT 1800
A V L R V V L G L L G E R Y I H L I N S 599
GCTGACTTCTCAGGTCAATATCCAGTCCATCTCGCTGTAAGGAAGGATGGAGAGCACTGT 1860
A D F S G Q Y P V H L A V R K D G E H C 619
CTTCGATTGCTAGTGGAGGCAGGAGCAAAGATTAACATGCCAGAGCAGAAAAGTGGCTGT 1920
L R L L V E A G A K I N M P E Q K S G C 639
ACAGCTCTACACCTGGCTGTGAGGGATAACCTGTAAAACTGGCCTGTAATCTTATTACT 1980
T A L H L A V R D N L L K L A C N L I T 659
GAGCTAAAGGCTGACGTGAATGTGTGTACCTACGGGGGAACAGTCCTCTTCACTTTGCT 2040
E L K A D V N V C T Y G G N S P L H F A 679
GCCAGTCAGGGCTCCCCACCTCTCTGCTCCATGCTGATAGCTGCAGGTGCTGATAAACGT 2100
A S Q G S P P L C S M L I A A G A D K R 699
ATGGAGAATGATGAGCCTCTTTTCCCTAAGCTCGTCTTCATCAGATGAAGATGATGAGGGT 2160
M E N D E P L F L S S S S S D E D D E G 719
GAGAAGAATGATGCTGATAGGCAGCTTCCCGATGACATGGTTCAGCAAGTTCCTTGATG 2220
E K N D A D R Q L P D D M V Q Q V H L M 739
TCAATATCATCCGAGAGGGGAACATTTAACCCACGCAAGAGACCTGCTGCAGGACACACA 2280
S I S S E R G T F N P R K R P A A G H T 759
CCATTTGACCTGGCTAAATGCCAAAAGTTAGAGATCTATTAGATGGCAGGAAAGGCCCC 2340
P F D L A K C Q K V R D L L D G R K G P 779
AAGCCAGTTCACCTCGCCCCAAGAGGACCAAAATAAGCACTGATGAAGTGAATCAAGGT 2400
K P S S L A P K R T K I S T D E V N Q G 799
TTGGACAGCGAGGTGCTCATGAACTGTGTGGTATTCTTATTGAAGACCATGTACCCTGG 2460
L D S E V L L M K L C G I L I E D H V P W 819
AGAGAGCTGGCAGAAAAACTGGGCATGCTCACACTCACTCATCTGTTCCAGGAGAGTGCC 2520
R E L A E K L G M L T L T H L F Q E S A 839
TCACCCTGCCAGAAGCTTCTCGAGAACTACCAGTTGAGTGGTGGTCCTGTGGACGGCCTG 2580
S P C Q K L L E N Y Q L S G G P V D G L 859

```

GTGGAGGCGCTGCAGTCTCTGGGTGTGAGTGAAGGAGTGAGACTTCTGACAGCCAGCCAA 2640
V E A L Q S L G V S E G V R L L T A S Q 879
CCCAGAGAGGACAAGCAGAGCTCAGATTCTACAGAGGACAGTGGCTTTGGGAGTCAGTCC 2700
P R E D K Q S S D S T E D S G F G S Q S 899
ATTGGAGAGGGGATGGGAAATCCTGAAGTGACCAATCACTTGAattatcaggatctattcat 2760
I G E G M G N P E V T N H - 913
tgaggacaacttggcactttatattgactttttttcaggtgcgaactggtggactgttcatc 2820
atctcaatgtggaaatacagcacgtcacaggaaggattgaaaagaatcgttttgaggtca 2880
aaagccctcactggtattacataacagtaagacatgacagaaatcagttgataactcaa 2940
ctctgaggtctggagtgtagctgtagctgtagtttttagtgcagacaattcagtgcttgta 3000
tctgaacagcaaataatccaactgagactttacagtaatgtaacagtaatgcaactgtcat 3060
ttccccagcacagctataaccttggaaatattgtttgtttgcagaatgttctatggatgtt 3120
tctagtaagctctctctcgatataatgcaacttgcaaaatgcaagatgtgatgaca 3180
atttaaatgactttgcttaattttcgatacatctagagatgtctatccaataactgtttg 3240
tttatttaatgtgtattttttggttgcagttatactttttgttcagaggttgagtaaataatgt 3300
ttgttttaataaaattattattttcgaaaaagaaaaaaaaaaaaaaaaaaaaa 3351

```

그림 1. 초어*nfkb*유전자의 핵산과 아미노산배열

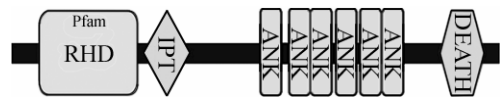
폴리아데닐산신호 aataaa와 attta배열, 번역개시코돈 ATG, 종결코돈 TGA는 강조체로 표시함

이 유전자에 의하여 암호화되는 단백질은 912개의 아미노산으로 되어있다. 3'-비번역배열에는 폴리아데닐산신호(aataaa)가 있고 mRNA배열의 불안정성을 보여주는 2개의 attta배열들이 있다. 선행연구[2]에 의하면 이 배열은 mRNA가 빠른 속도로 분해되게 하는 기능을 수행한다. Blastp분석결과 번역된 배열은 이미 밝혀진 줄말고기의 *NFKB*단백질들과 상동성이 매우 높았다.(90.7%)

그리하여 이 유전자의 이름을 초어학명에 있는 속명과 종명의 첫 글자를 따서 *Cinfb*로 명명하였다.

SMART와 TMHMM프로그램으로 도메인구조를 분석한데 의하면 *Cinfb*유전자에 의하여 암호화되는 단백질은 1개의 RHD도메인과 1개의 IPT도메인, 6개의 ANK도메인, 1개의 DEATH도메인을 가지고있다.(그림 2)

핵인자단백질은 여러가지 신호통로들에서 매우 중요한 기능을 수행하는 신호매개단백질이며 세포막에서 생겨난 신호를 넘겨받은 다음 핵인자억제단백질과 분리되면서 활성화되어 세포핵속에 들어가 전사조절인자로서 작용한다.[3]

그림 2. 초어의 핵인자(*Cinfb*)의 도메인구조

초어*nfkb*유전자의 배열과 구조에 대한 연구결과들은 물고기류에서도 포유동물과 마찬가지로 *nfkb*유전자가 신호통로를 통하여 선천성면역반응을 일으키는 기능을 수행할수 있다는것을 보여준다.

맺 는 말

1) 초어*nfkb* cDNA의 전배열크기는 3 351bp이고 5'-비번역배열과 열린읽기틀, 3'-비번역배열의 크기는 각각 3, 2 739, 609bp이며 912개의 아미노산을 암호화한다.

2) 초어*nfkb*유전자에 의하여 암호화되는 단백질은 1개의 RHD도메인과 1개의 IPT도메인, 6개의 ANK도메인, 1개의 DEATH도메인을 가지고있다.

참 고 문 헌

- [1] J. R. Sambrook; Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 11~25, 2001.
- [2] A. E. Kel et al.; Nucleic Acids Res., 31, 3576, 2003.
- [3] S. Holger et al.; Diabetes, 55, 2993, 2016.

주체107(2018)년 1월 5일 원고접수

cDNA Sequence Characteristics of *nfkB* Gene from Grass Carp, *Ctenopharyngodon idella*

An Jun Hyok, Jang Song Hun

The full length of the grass carp *nfkB* cDNA is 3 351 nucleotides(nt). The 5'-untranslated region(UTR) is 3nt and open reading frame is 2 739nt. The 3'-UTR is 609nt including a poly A tail. The 912 amino acid polypeptide encoded by this gene. The protein encoding by grass carp *nfkB* gene has one RHD domain, one IPT domain, six ANK domain and one DEATH domain.

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