

초어의 핵인자유전자(*Cinfkb*)의 게놈구조와 도메인구조의 분석

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 발전시켜나갈수 있습니다.》(《김정일선집》 증보판 제10권 485페이지)

유전자의 분자구조특성을 밝히는것은 해당 유전자의 기능을 깊이있게 해명하고 실천에 도입하는데 필요한 기초자료를 마련하는데서 큰 의의를 가진다.[1, 2]

우리는 초어의 선천성면역조절에서 중요한 역할을 하는 *Cinfkb*의 게놈 및 도메인구조에 대한 연구를 하였다.

재료와 방법

재료로는 초어핵인자유전자(*Cinfkb*)의 게놈배열을 리용하였다.

1) *Cinfkb*의 게놈구조분석

*Cinfkb*의 게놈구조는 이 유전자의 게놈배열과 cDNA배열을 NCBI웹사이트의 Spidey기능을 리용하여 엑손과 인트론을 갈라내는 방법으로 분석하였다.

*Cinfkb*의 게놈배열은 초어의 전게놈배열자료기지에서 얻어졌으며 cDNA배열은 우리가 분리한 배열을 리용하였다.

2) *Cinfkb*에 의하여 암호화되는 단백질의 도메인구조분석

*Cinfkb*의 cDNA배열을 ExPASy프로그램을 리용하여 아미노산배열로 넘겨 이 유전자의 가상적인 단백질배열을 얻어냈다. 다음 SMART 및 TMHMM프로그램으로 이 단백질배열을 처리하여 도메인구조를 분석하였다.[1]

결과 및 논의

1) *Cinfkb*의 게놈구조

*Cinfkb*의 게놈배열과 cDNA배열을 NCBI웹사이트의 Spidey프로그램으로 분석하면 이 유전자가 22개의 엑손과 21개의 인트론을 가지고있다는것을 알수 있다.(표 1, 2)

표 1에서 보는바와 같이 엑손의 크기범위는 41~221bp범위에 있으며 엑손 7과 엑손 18을 제외한 나머지엑손들에서 게놈배열과 cDNA배열이 일치된다. 엑손 7과 엑손 18에서 게놈배열과 cDNA배열의 동일성은 각각 99.4, 99.5%로서 100%에 이르지 못하는 원인은 이미 등록된 게놈배열들의 배열분석상오유라고 본다.

표 1. *Cinfbk*의 게놈에서 엑손들의 위치와 크기

엑손번호	게놈배열위치/bp	cDNA배열위치/bp	엑손크기/bp	동일성/%
엑손 1	201-242	16-57	42	100.0
엑손 2	815-893	58-136	79	100.0
엑손 3	1 026-1 066	137-177	41	100.0
엑손 4	1 718-1 816	178-276	99	100.0
엑손 5	1 900-2 051	277-428	152	100.0
엑손 6	2 228-2 340	429-541	113	100.0
엑손 7	2 464-2 622	542-700	159	99.4
엑손 8	4 033-4 137	701-805	105	100.0
엑손 9	4 220-4 308	806-894	89	100.0
엑손 10	4 424-4 562	895-1 033	139	100.0
엑손 11	5 732-5 866	1 034-1 168	135	100.0
엑손 12	5 969-6 145	1 169-1 345	177	100.0
엑손 13	6 240-6 381	1 346-1 487	142	100.0
엑손 14	6 676-6 790	1 488-1 602	115	100.0
엑손 15	6 950-7 157	1 603-1 810	208	100.0
엑손 16	7 273-7 442	1 811-1 980	170	100.0
엑손 17	9 081-9 183	1 981-2 083	103	100.0
엑손 18	9 531-9 751	2 084-2 304	221	99.5
엑손 19	9 860-9 941	2 305-2 386	82	100.0
엑손 20	10 025-10 188	2 387-2 550	164	100.0
엑손 21	11 151-11 262	2 551-2 662	112	100.0
엑손 22	11 364-11 440	2 663-2 739	77	100.0

표 2에서 보는바와 같이 인트론들의 크기는 82~1 638bp범위에 있는데 대부분이 1 000bp 이하이며 1 000bp이상의 크기를 가지는 인트론은 인트론 7과 10, 16밖에 없다는것을 알수 있다.

표 2. *Cinfbk*게놈에서 인트론들의 위치와 크기

인트론번호	게놈배열위치 /bp	인트론크기 /bp	인트론번호	게놈배열위치 /bp	인트론크기 /bp
인트론 1	243-814	572	인트론 12	6 146-6 239	94
인트론 2	894-1 025	132	인트론 13	6 382-6 675	294
인트론 3	1 067-1 717	651	인트론 14	6 791-6 949	159
인트론 4	1 817-1 899	83	인트론 15	7 158-7 272	115
인트론 5	2 052-2 227	176	인트론 16	7 443-9 080	1 638
인트론 6	2 341-2 463	123	인트론 17	9 184-9 530	347
인트론 7	2 623-4 032	1 410	인트론 18	9 752-9 859	108
인트론 8	4 138-4 219	82	인트론 19	9 942-10 024	83
인트론 9	4 309-4 423	115	인트론 20	10 187-11 150	964
인트론 10	4 563-5 731	1 169	인트론 21	11 263-11 363	101
인트론 11	5 867-5 968	102			

*Cinfbk*의 게놈구조를 도식화하였는데(그림 1) 진한 색으로 표시한것은 엑손이며 그 너비는 엑손의 크기를 표시한다. 연한 색으로 표시한것은 인트론으로서 그 너비는 매개 인트론의 크기를 표시한다. 엑손은 *Cinfbk* 게놈배렬의 201bp 위치에서 시작되어 11 440bp위치에서 끝난다.

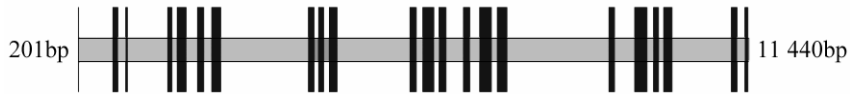


그림 1. *Cinfbk*의 게놈구조

2) *Cinfbk*에 의하여 암호화되는 단백질의 도메인구조

Expassy프로그램을 리용하여 *Cinfbk*의 cDNA배렬로부터 가상적인 초어NFKB배렬을 얻어내고 SMART 및 TMHMM프로그램으로 단백질구조를 분석한 결과는 표 3, 그림 2와 같다.

표 3에서 보는바와 같이 가상적인 초어NFKB단백질에는 1개의 RHD도메인과 1개의 IPT도메인, 4개의 low complexity모티프배렬, 6개의 ANK모티프배렬, 1개의 DEATH도메인이 존재한다는것을 알수 있다.

표 3. 초어NFKB단백질에서 도메인들의 위치와 크기

도메인이름	시작위치/aa	마감위치/aa	도메인크기/aa
Pfam:RHD	51	233	183
IPT	240	340	101
low complexity	366	392	27
low complexity	399	419	21
low complexity	434	447	14
ANK	493	522	30
ANK	532	561	30
ANK	565	594	30
ANK	603	632	30
ANK	637	667	31
ANK	671	700	30
low complexity	709	720	12
DEATH	794	879	86

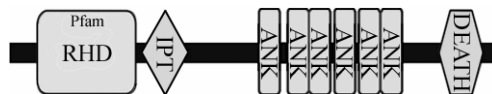


그림 2. 초어NFKB단백질의 도메인구조

핵인자단백질은 여러가지 면역관련신호통로들에서 매우 중요한 기능을 수행하는 신호매개단백질이며 세포막에서 생겨난 신호를 넘겨받은 다음 핵인자억제단백질과 분리되면서 활성화되어 세포핵속에 들어가 전사조절인자로서 작용한다.[2]

*Cinfbk*의 게놈구조와 도메인구조에 대한 연구결과들은 물고기류에서도 포유동물과 마찬가지로 *Cinfbk*가 신호통로를 통하여 선천성면역반응을 일으키는 기능을 수행할수 있다는

것을 보여준다.

맺는 말

- 1) *Cinfkb*는 22개의 엑손과 21개의 인트론으로 이루어졌다.
- 2) *Cinfkb*에 의하여 암호화되는 단백질은 1개의 RHD도메인과 1개의 IPT도메인, 6개의 ANK도메인, 1개의 DEATH도메인을 가지고있다.

참고 문헌

- [1] A. E. Kel et al.; Nucleic Acids Res., 31, 3576, 2003.
[2] S. Holger et al.; Diabetes, 55, 2993, 2016.

주체107(2018)년 4월 5일 원고접수

Genomic Organization and Domain Architecture of *Cinfkb* Gene from Grass Carp, *Ctenopharyngodon idella*

An Jun Hyok, Jang Song Hun

The grass carp *nfbk* gene consists of 22 exons and 21 introns. The protein encoding by this gene consists of 904 amino acids. The protein encoding by *Cinfkb* gene has one RHD domain, one IPT domain, six ANK domains and one DEATH domain.

Key words: grass carp, *Ctenopharyngodon idella*, *Cinfkb*, genomic organization, domain architecture