

치매환자에서 transposable elements에 의한 유전자 발현조절의 생물정보 분석

김대수¹ · 허재원² · 하홍석² · 김태홍² · 조운종¹ · 김희수^{1,2*}¹부산대학교 자연과학대학 생물정보학협동과정²부산대학교 자연과학대학 생명과학부

Received August 8, 2006 / Accepted November 13, 2006

Bioinformatics Analysis of Gene Expression Regulation by Transposable Elements in Dementia Patients. Dae-Soo Kim¹, Jae-Won Huh², Hong-Seok Ha², Tae-Hong Kim², Un-Jong Jo¹ and Heui-Soo Kim^{1,2*}. ¹PBBRC, Interdisciplinary Research Program of Bioinformatics, College of Natural Sciences, Pusan National University, Busan 609-735, Korea, ²Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea – Dementia is a progressive disease of increasing the dysfunction of intellectual and physical ability. In the aging society, many families are suffering from the caring the patients who are diagnosed with dementia. However, dementia is a complex disease affected by the genetic and environmental agents. In the present study, we investigated the transposable elements in relation to dementia. From the analysis of dementia EST (expressed sequence tag) sequences, we found dementia candidate genes, and analyzed expression profiles and repeat elements using bioinformatics tools. This analysis showed that 98 genes were affected in their mRNA sequences by transposable elements expression. Their expressions were affected by the integration of different transposable elements (SINE, LINE, LTR, DNA) during the primate evolution. We believe that our work will be of significant interest to genome scientists, and may help them gain insight into implication of transposable elements expression in dementia.

Key words – Dementia, EST, bioinformatics, transposable elements

서 론

과학기술과 의료기술의 발전은 인류의 수명연장을 급속히 촉진시켰고, 이로 인한 노령인구의 증가는 새로운 사회 문제인 치매에 많은 관심을 불러일으켰다. 고령화 사회가 도래하면서 노인성 치매환자가 매년 급격히 늘고 있으며 매년 수만 명이 노인성치매에 걸려 본인뿐 아니라 가족 까지도 많은 고통에 시달리고 있다. 특히 노인성 치매의 경우는 환자 본인의 문제가 아니라 젊은 노동력을 환자의 보호자로 필요함으로 국가적 노동력의 손실로 이어지고 있다. 노인질환인 치매의 경우는 한국뿐만 아니라 전세계적으로도 경제적, 사회적 문제로 대두되면서 인류가 급히 해결해야 할 범 지구적인 질병으로 규정되고 있다[15]. 노인성 질환인 치매(Dementia)는 노년에 나타나는 정신병적 증상의 가장 흔한 원인 질환으로 만성적이고 서서히 악화되는 진행성이며, 기억력, 사고력, 학습능력 및 판단력 등의 손상을 포함한 인지기능의 장애이다 [1,2].

치매환자는 크게 두 가지의 경우로 나눌 수 있는데 하나는 뇌 혈관성 질환과 또 하나는 신경세포 자체가 고사되어 생기는 알츠하이머병(Alzheimer's disease)로 나눌 수 있다 [20]. 또한 Alzheimer의 경우 APP (amyloid precursor

protein), PSEN (Presenilin), APOE (Apolipoprotein E)등의 유전자가 뇌의 신경세포를 고사시키는 원인으로 규명되어 있고, 현재 많은 연구가 진행되고 있다[16]. APP유전자의 경우 amyloid 전구 단백질이 plaque 형태로 뇌의 뉴런에 α -peptide 와 β -peptide가 침적되어 시냅스상실이나 신경원섬 유농축 형성에 의해서 발생하는 신경퇴행성 현상이다 [16,18]. 이러한 질병원인 단백질인 α 와 β 는 transmembrane protein인 APP 단백질이 β -secretase 와 γ -secretase에 의해 절단되어 생성되며, 정상인 경우 주로 $A\beta$ -40이 생성되지만 Alzheimer의 경우 $A\beta$ -42가 더욱 침착되어 Alzheimer의 원인이 되는 것으로 알려져 있다[18]. PSEN (Presenilin) 유전자의 이상으로 발생하는 Alzheimer의 경우는 조기유발성 Alzheimer이다[3]. PSEN과 APP유전자와는 서로 상호작용을 하고 있는 것으로 추정되고 있으며 PSEN-2유전자의 돌연변이가 발생하면 APP의 γ -secretase에 의한 절단을 증가시켜 $A\beta$ -42의 신경독성섬유를 침적시키는 원인이 된다[5,25]. 또한 APOE는 유전자 형에 따라 E2, E3, E4라는 세가지 유형으로 구분되며, 이중 E4 유전자형을 가진 경우 치매의 발병 위험성은 높다[19]. apoE4가 Alzheimer의 주 위험 요인인 이 유는 β -secretase 단백질과 상호 작용함으로써 amyloid 가 마치 머리카락 같은 모양을 이루고, 손상된 뇌신경 세포와 함께 뉴런 plaque를 만들어 섬유 모양이 되어 신경 퇴화를 촉진시키기 때문이다[6,26,27]. 이러한 유전학적인 많은 연구들은 Alzheimer의 질병원인과 질병치료를 밝히기 위해 많

*Corresponding author

Tel : +82-51-510-2259, Fax : +82-51-581-2962

E-mail : khs307@pusan.ac.kr

은 연구가 이루어지고 있다. 하지만 아직까지 Alzheimer를 발생하는 원인을 몇몇 유전자의 돌연변이에 발생되고 있다는 연구결과만 있을 뿐 확실한 원인 규명이 되지 않고 있다.

인간의 지놈은 transposable elements가 45%정도를 차지하고 있으며 이들은 크게 4개의 범주에 의해서 나뉜다[27]. 자체적인 유전자의 발현능력을 가지지 못하는 SINE과 DNA family와 자체적인 유전자의 발현능력을 가지는 LINE과 LTR family로 분류된다[15,19]. 대부분의 transposable elements들은 유전자의 발현에 영향을 줄 수 있기 때문에 진화적인 선택압에 의해 내부에 많은 돌연변이가 축적 되어 따라 완전한 구조를 가지는 family들이 많지 않다[17,22]. 하지만 최근의 연구들에서 이들 transposable elements들이 유전자의 발현을 조절하여 질병을 유발하거나 유전자의 다양성에 많은 기여를 한다는 연구결과들이 밝혀지고 있다[4,9,14]. 일반적으로 'Junk DNA'로 알려진 transposable elements들이 인류의 진화과정 동안에 유전자의 다양성에도 많은 기여를 하고 있으며 인간의 질병에도 밀접한 관련이 있다는 연구들의 결과를 토대로 우리는 현대의 커다란 문제로 떠오르는 노인성질환인 치매와의 상관관계를 생물정보학적인 분석방법을 이용하여 *in silico* 분석을 하여 치매의 원인을 규명하고자 공개된 유전자 데이터베이스의 EST 서열을 이용하여 분석하였다.

재료 및 방법

치매환자에서 획득한 EST (Expressed Sequence Tags) 분류

치매와 transposable elements와의 상관관계를 분석하기 위하여 우선 공개된 유전자 데이터베이스인 NCBI의 dbEST 서열을 이용하여 치매환자의 뇌 조직에서 클로닝하여 만들어진 유전자 조각인 EST (expressed sequence tags)를 분류하였다. 치매에 관련된 EST를 분리하는 기준으로는 NCBI의 genbank파일에서 조직정보가 Alzheimer cortex 이거나 Alzheimer brain 정보를 가지는 EST를 정보를 이용하였으며 조직정보가 명확하지 않은 EST서열은 분석하는데 사용하지 않았다. 결과적으로 치매환자의 뇌 조직에서 만들어진 11,381개의 EST를 분류하였다.

EST 클러스터링

치매환자의 뇌에서 유래한 11,381개의 EST를 클러스터링하기 위하여 우선 전처리 작업으로 각 EST서열에 있는 벡터 서열을 제거하였다. 벡터서열을 제거하기 위한 방법으로 RepeatMasker (<http://repeatmasker.genome.washington.edu>) 프로그램을 이용하였으며 사용한 벡터서열은 NCIB (<http://www.ncbi.nlm.nih.gov/>) 데이터베이스에서 UniVec 서열을 이용하였다. 전처리 작업을 수행한 EST를 클러스터

링하기 위한 방법으로는 슈퍼바이저 방법을 이용하여 잘 정리된 인간의 지놈에 EST서열을 맵핑하는 방법을 채택하였다. EST 서열을 정렬하는데 MegaBlast[28] 프로그램을 이용하여 서열의 유사도가 97% 이상 되게 하여 EST서열을 클러스터링 하였다. 또한 클러스터링된 EST의 유전자의 구조를 분석하기 위하여 SIM4[7] 프로그램을 이용하여 지놈상에서 엑손과 인트론의 구조를 분석하였다. 또한 인간의 지놈에 유전자의 위치를 찾기 위하여 NCBI의 RefSeq서열을 이용하였으며 이때 사용한 프로그램은 SIM4프로그램을 이용하였다. 우리는 EST가 클러스터링된 영역이 인간 유전자 영역에 포함되는 클러스터와 유전자 영역이 아닌 영역에 정렬된 EST 클러스터를 분류하였으며 우리는 유전자 영역이 아닌 다른 영역에서 정렬된 EST 클러스터는 분석하는데 사용하지 않았다.

치매 전사체에서의 transposable elements 분석

치매 환자의 뇌조직에서 만들어진 EST에서의 transposable elements의 발현을 분석하기 위하여 우리는 EST클러스터가 포함된 유전자의 영역의 앞뒤 서열을 5kb를 확장하여 유전자 서열을 얻었다. Transposable elements가 유전자의 앞뒤에서 유전자의 발현에 영향을 주어 발생할 수 있기 때문에 우리는 유전자의 앞뒤로 5kb를 확장하여 분석하는데 사용하였다. 획득한 유전자 서열을 RepeatMasker 프로그램을 이용하여 유전자서열 내에 존재하는 모든 transposable elements를 찾아 내었다. 유전자 내에서 분포하고 있는 transposable elements 가 치매에 걸린 환자에서 발현되는지를 분석하기 위하여 치매 EST서열의 정렬된 결과와 transposable elements의 영역과의 비교를 통해서 transposable elements 가 치매 EST에서 부분적으로 발현되는 결과만을 분류하였다. 이때 우리가 사용한 기준은 EST가 적어도 인트론에 의해서 2개의 엑손으로 분리가 되어야 하며 또한 엑손과 인트론으로 나뉠 때 스플라이싱 사이트인 (AG-GT)를 가지는 것만을 분석에 사용하였다.

Dual coding 유전자에서 발현하는 transposable elements

일반적인 유전자들 중에서 인류의 진화과정 동안 유전자의 인트론에 삽입된 transposable elements가 유전자가 전사체를 만드는 과정에서 splicing out되지 않고 유전자의 일부 분처럼 발현되어 transposable elements에 의해 다양한 전사체를 만들어내는 유전자를 찾기 위해서 우리는 NCBI의 RefSeq서열에서 2개이상 다양한 전사체를 가지는 유전자를 얻었다. 이렇게 얻어진 유전자를 인간 지놈서열에 정렬하기 위하여 SIM4프로그램을 이용하였으며 이때 사용한 기준은 97%이상의 서열의 유사성을 가지도록 하였다. SIM4결과를 이용하여 transposable elements의 일부분의 서열이 유전자의 엑손처럼 발현되는 것을 찾기 위하여 유전자서열을 RepeatMasker를 실행하여 유전자 위치정보와 transposable

elements의 위치정보를 이용하여 transposable elements의 일부분이 엑손으로 발현되는 유전자의 정보를 얻었다.

결과 및 고찰

Dual coding 유전자의 다양성을 만들어 내는 transposable elements

우리는 인간지놈에 45%이상을 차지하고 있는 transposable elements가 인류의 진화과정 동안에 얼마나 많은 유전자에서 유전자의 다양성을 만들어내고 있는지를 분석하였다[13]. Table 1의 데이터는 transposable elements의 단편이 유전자의 전사체에 삽입되어 유전자서열의 일부처럼 발현되고 있는 데이터들 중에 인간의 치매에 관련하여 연구 중이거나 많은 연구가 시행된 유전자들이다. 결과적으로 인류의 진화과정 동안에 많은 transposable elements가 유전자의 인트론 영역에 삽입되어있으면서 전사체를 만드는 과정에서 splicing out되지 못하고 transposable elements에 존재하는 splicing site (AG-GT)를 인식하여 유전자서열화 되었으며 또한 유전자의 UTR (untranslated region)이 아닌 CDS (coding sequence)에서 유전자의 일부분이 되어 발현되고 있다. 하나의 유전자에서 다양한 형태의 전사체를 만들어내는 기작으로 잘 알려진 alternative splicing의 경우 유전자의 엑손의 조합으로 인해서 다양한 형태의 전사체를 만든다[8]. 유전자의 인트론 영역에 많이 분포하고 있는 transposable elements들이 때로는 alternative splicing에 의해서 유전자의 엑손화의 과정을 거쳐 유전자의 다양성을 야기시키는 연구 결과들이 최근에 보고되었다[24]. Table 1의 결과 또한 alternative splicing의 여러 기작들 중에서 인트론의 exonization을 통하여 유전자의 다양성을 제공하는 경우라 할 수 있다. 소수의 유전자에서 transposable elements가 프로테인으로 만들어지는 코딩영역에서 발현되지는 않은 결과를 얻었다. 하지만 transposable elements들이 유전자와는 상반되는 진화적인 선택압을 받고 있기 때문에 상대적으로 많은 돌연변

이를 발생했을 것이라 생각되고 돌연변이가 많이 발생했을 경우에는 기존의 어떠한 transposable elements들과의 서열의 유사성이 많이 떨어지기 때문에 일률적인 서열정렬 방법으로는 유전자 서열화된 transposable elements 들을 찾지 못했을 것이라 생각된다. 더 많은 유전자에서 오랜 시간 동안 진화를 거치면서 transposable elements에 의한 유전자의 다양성이 제공되었을 것이라 생각되고 transposable elements는 유전자의 다양성을 제공하는 중요한 요소로서 작용하고 있을 것으로 추정된다. 이러한 유전자의 다양성을 야기시키는 중요한 요소로서 작용하고 있는 transposable elements를 진화적인 관점에서 분석함으로써 transposable elements가 어느 시기에 인간의 유전자에서 발현되었으며 유전자의 다양성을 획득하였는지를 분석한다면 인류의 진화를 이해 할 수 있을 것으로 추측된다.

치매 환자의 전사체에서 발현하는 transposable elements

치매환자의 뇌조직을 이용하여 만들어진 EST (expressed sequence tag)서열은 치매 환자의 뇌조직에서 발현하는 유전자들의 단편이기 때문에 유전자의 비정상적인 발현에 의한 치매의 발병원인을 유전자에서 발생하는 돌연변이가 아닌 transposable elements의 이상 발현으로 인한 치매의 원인을 규명하고자 EST 서열의 분석을 통해, transposable elements와 치매와의 상관관계를 분석하였다. 치매를 발생시키는 여러 요인들이 존재하지만, 이전 치매연구에서 분석하지 않았던 transposable elements에서 치매의 원인을 규명하기 위하여 분석한 결과 총 98개의 유전자에서 맵핑된 치매 EST에서 transposable elements가 발현되고 있음을 확인할 수 있었다. 98개의 유전자에 맵핑된 치매 EST에서 발현되고 있는 transposable elements들이 치매의 직접적인 원인이 되고 있는지는 실험적인 증명이 필요하지만 생물정보학적인 분석방법을 이용하여 분석하여 98개의 치매관련 후보유전자를 예측하였다. 또한 98개의 치매관련 후보 유전자에서 발현되고 있는 transposable elements의 대부분이 인트론 영역에 삽입되어

Table 1. The list of Dual coding genes derived by the transposable elements

Gene Name	Chr	Gene description	Total Transcript	TEs Fusion		Fusion TEs family		
				Transcript NO.	Fusion TEs	Family	TEs NO.	
APBB1	11	Amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	2	2	LTR50	LTR/ERV L	2	
BAX	19	BCL2-associated X protein	6	1	AluSx	SINE/Alu	1	
DISC1	1	Disrupted in schizophrenia 1	4	1	MIR	SINE/MIR	1	
GABRB2	5	Gamma-aminobutyric acid (GABA) A receptor, beta 2	2	1	L3	LINE/CR1	1	
GRK2	6	Glutamate receptor, ionotropic, kainate 2	2	2	L2	LINE/L2	2	
MOBP	3	Myelin-associated oligodendrocyte basic protein	3	2	LTR33A	LTR/ERV L	2	
MOG	6	Myelin oligodendrocyte glycoprotein	9	2	AluJb	SINE/Alu	2	
NFS1	20	NFS1 nitrogen fixation 1 (S. cerevisiae)	2	1	AluY	SINE/Alu	1	
PLAUR	19	Plasminogen activator, urokinase receptor	3	1	MLT1G	LTR/MaLR	1	
POLR2J2	7	DNA directed RNA polymerase II polypeptide J-related gene	3	2	MSTA	LTR/MaLR	2	
PSEN1	14	Presenilin 1 (Alzheimer disease 3)	3	1	MERS2A	LTR/ERV 1	1	
SLC3A2	11	Solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	6	4	FLAM_C	AluSq SINE/Alu	8	
SLCO1A2	12	Solute carrier organic anion transporter family, member 1A2	3	1	MER70B	LTR/ERV L	1	
UBE2J2	1	Ubiquitin-conjugating enzyme E2, J2 (UBC6 homolog, yeast)	5	1	AluSx	SINE/Alu	1	
UBE2L3	22	Ubiquitin-conjugating enzyme E2L 3	2	1	L2	LINE/L2	1	

있었으면서 전사체가 만들어지는 과정에서 exonization되어 유전자처럼 발현되는 데이터가 대부분이었다. 치매환자의 EST에서 발현하고 있는 transposable elements를 분석하기 위하여 transposable elements를 크게 5개의 범주로 나눠서 치매에 많은 연관성을 가지는 transposable elements를 분석하였다. Fig. 1의 결과에서 나타내는 것처럼 SINE이 가장 많은 수가 발현되고 있음을 알 수 있으며 그 다음으로 LINE, DNA, LTR 순서로 발현이 많이 되고 있음을 확인할 수 있다. 인간의 지놈의 45%정도를 transposable elements를 분포하고 있으며 (Fig. 2) transposable elements는 GC contents에 따라 서로 상반되는 영역에 분포하는 경향을 가진다. SINE의 경우는 LINE과 LTR과는 다른 GC contents가 높은 영역에 많이 분포되어있으며 유전자 또한 GC contents가 높은 영역에 많이 분포한다[21]. GC contents의 분포에 따른

transposable elements의 분포의 차이 때문에 SINE이 유전자 발현에 좀더 많은 영향을 주었을 것이라 추측된다.

치매 전사체에서 발현하고 있는 transposable elements 발현 패턴

치매 환자의 EST에서 발현되고 있는 transposable elements들의 각각의 family를 인간의 지놈에 삽입된 시기에 따라 각 subfamily로 분류하여 각 subfamily별 발현 정도를 분석하였다. Table 2의 데이터는 인간의 진화과정 동안에 삽입된 시기가 가장 많이 분석된 SINE의 family의 데이터를 나타낸다. Table 2의 데이터에서 보는 것처럼 인간 지놈에 삽입된 시기가 가장 오래된 SINE이 가장 많이 발현되고 있음을 알 수 있다. 오래 전에 삽입된 SINE의 경우 가장 최근에 삽입된 SINE보다 많은 진화적인 선택압에 따라 많은 돌

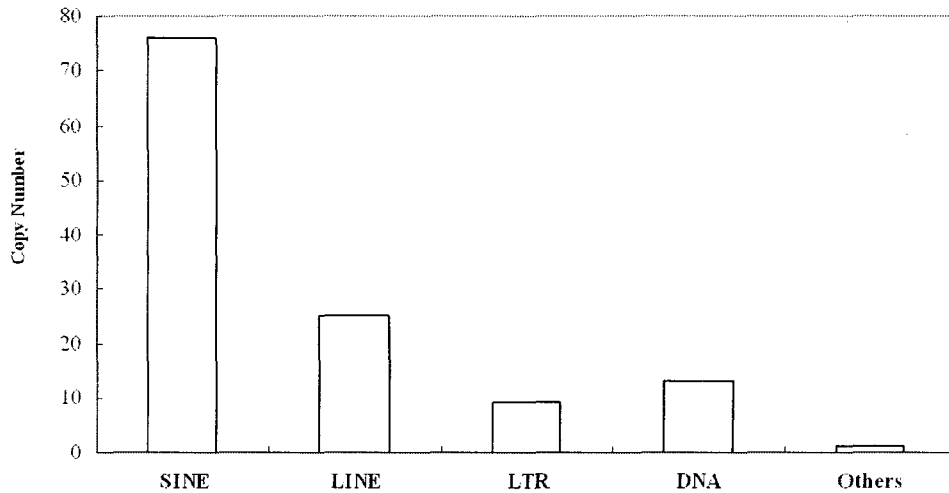


Fig. 1. Expression profile of transposable elements in dementia patient's EST.

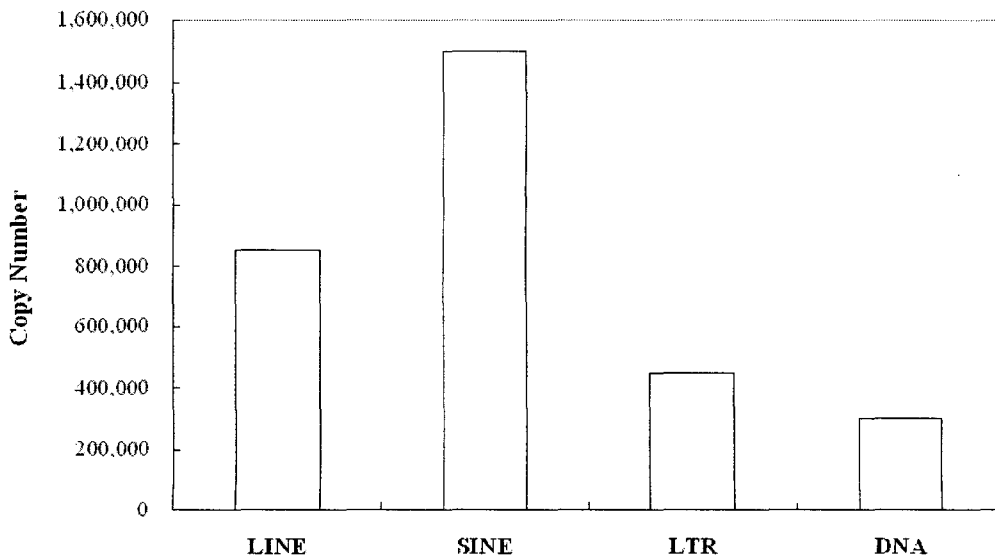


Fig. 2. Distribution of transposable elements in human genome.

Table 2. Distributions of SINE subfamilies within SINE-containing dementia patient's EST

SINE Subfamily	Age (Million years)	Distribution of TEs	
		Occurrences	Percent
Alu	112	0	0.00
AluJ	81	12	15.79
AluS	31-48	21	27.63
AluY	19	9	11.84
FRAM	-	3	3.95
MIR	-	31	40.79

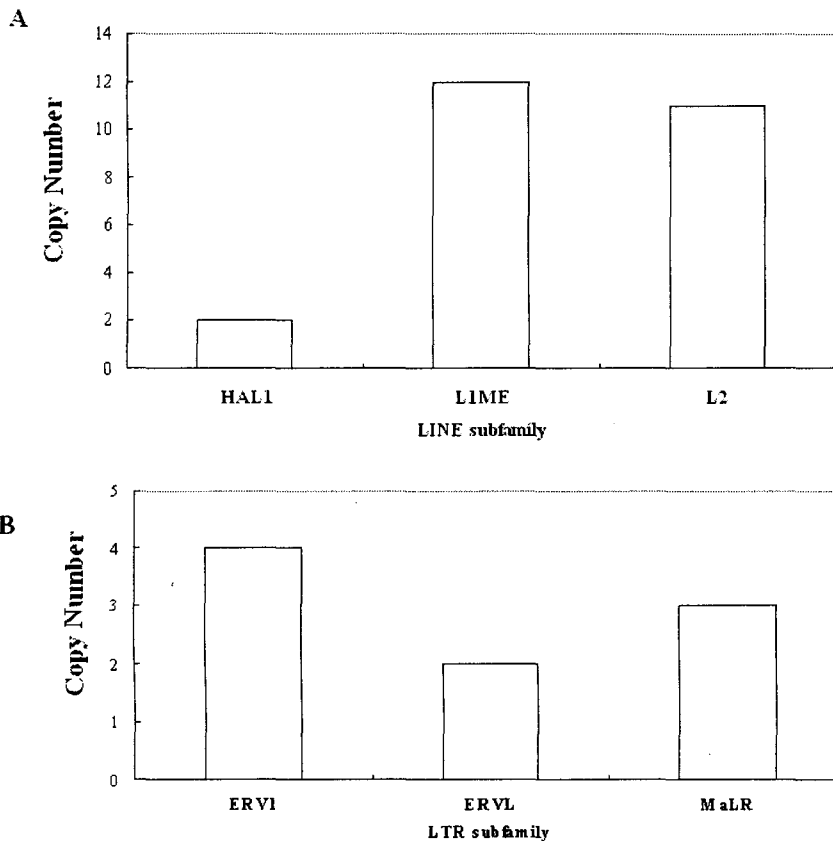


Fig. 3. Distributions of LINE and LTR subfamilies within transposable elements containing dementia patient's EST.

연변이들이 쌓이게 됨으로써 transposable elements의 원래의 성질을 잃게 되고 지능화 되어 가는 과정에서 많은 돌연변이의 축적으로 인해 splicing site를 얻게 되어 오래된 transposable elements 들이 최근에 삽입된 transposable elements들보다 더 많이 발현되었을 것으로 추측된다. LINE과 LTR의 경우 또한 SINE과 비슷한 양상을 보여줌을 확인하였다(Fig. 3A, 3B) LINE과 LTR은 유전자의 발현에 영향을 미치는 프로모터들을 가지고 있다. 그렇기 때문에 다른 transposable elements들에 비하여 많은 돌연변이가 발생하여 대부분 완전한 구조를 가지지 못하고 변형된 구조를 많이 가진다. 하지만 이들이 유전자의 앞쪽에 삽입되어있음으로써 인간 유전자의 발현을 조절하는 경우가 최근 많은 연구들을 통

하여 밝혀지고 있다[11,12]. 치매환자에서 만들어진 EST서열의 양이 적었기 때문에 상대적으로 LINE과 LTR의 프로모터에 의해서 발현이 조절되는 데이터를 얻지 못하였을 것으로 추측된다. 유전자의 발현에 영향을 주는 transposable elements의 분석을 통하여 인류의 진화과정을 해석한다면 인류의 질병 퇴치에 크게 기여할 것이다.

요 약

노년에 나타나는 정신병적 증상의 가장 흔한 원인 질환으로 만성적이고 서서히 악화되는 진행성이며, 기억력, 사고력, 학습능력 및 판단력 등의 손상을 포함한 인지기능의 장애이

다. 고령화 사회가 도래하면서 노인성 치매환자가 매년 급격히 늘고 있으며 매년 수만 명이 노인성치매에 걸려 본인뿐 아니라 가족 까지도 많은 고통에 시달리고 있다. 특히 노인성 치매의 경우는 환자 본인의 문제가 아니라 젊은 노동력을 환자의 보호자로 필요함으로 국가적 노동력의 손실로 이어지고 있다. 현재 연구에서 우리는 노인성 치매와 transposable elements와의 상호관계를 밝히기 위하여 공개된 유전자 데이터베이스에서 EST (expressed sequence tags)를 이용하여 생물정보학적 인 분석방법과 프로그램을 이용하여 치매의 원인으로 추정되는 후보유전자들을 찾아내었다. 이러한 분석을 통하여 치매환자에서 transposable elements의 발현으로 인해 유전자의 발현에 변화를 가지는 98개의 후보 유전자를 찾아내었다. 노인성질환인 치매와 transposable elements의 분석방법을 이용하면 치매의 원인을 규명하는데 많은 도움이 될 것이다.

감사의 글

이 논문은 2005년도 국립대학구조개혁 지원 사업비에 의하여 연구되었음.

참고 문헌

- Alfredo, A. S. 1988. Assessing vision in patients with Alzheimer's disease. *West. J. Med.* **148**,693.
- Ball, M.J., M. Fisman, V. Hachinski, W. Blume, A. Fox, V. A. Kral, A. J. Kirshen, H. Fox and H. Merskey. 1985. A new definition of Alzheimer's disease: a hippocampal dementia. *Lancet* **8419**, 14-16.
- Borchelt, D.R., G. Thinakaran, C. B. Eckman, M. K. Lee, F. Davenport, T. Ratovitsky, C. M. Prada, G. Kim, S. Seekins and D. Yager. 1996. Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio in vitro and in vivo. *Neuron.* **17**, 1005-1013.
- Christensen, T. 2005. Association of human endogenous retroviruses with multiple sclerosis and possible interaction with herpes viruses. *Rev. Med. Virol.* **15**, 179-211.
- Deng, G., C. J. Pike and C. W. Cotman. 1996. Alzheimer-associated presenilin-2 confers increased sensitivity to apoptosis in PC12 cells. *FEBS Lett.* **397**, 50-54.
- Farrer, L.A., L. A. Cupples, J. L. Haines, B. Hyman, W. A. Kukull, R. Mayeux, R. H. Myers, M. A. Pericak-Vance, N. Risch and C. M. Van Duijn. 1997. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *J. Am. Med. Assoc.* **278**, 1349-1356.
- Florea, L., G. Hartzell, Z. Zhang, G. M. Rubin and W. Miller. 1998. A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* **8**, 967-974.
- Graveley, B.R. 2001. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* **17**,100-107.
- Halling, K.C., C. R. Lazzaro, R. Honchel, J. A. Bufill, S. M. Powell, C. A. S. Arndt and N. M. Lindor. 1999. Hereditary desmoid disease in a family with a germline *Alu I* repeat mutation of the APC gene. *Hum. Hered.* **49**, 97 - 102.
- Kazazian, Jr.HH. 2004. Mobile elements: drivers of genome evolution. *Science*, **303**, 1626-1632.
- Kim, D.S., T. H. Kim, J. W. Huh, I. C. Kim, S. W. Kim, H. S. Park and H. S. Kim. 2006. LINE FUSION GENES: a database of LINE expression in human genes. *BMC Genomics.* 2006, 7,139.
- Kim, T.H., Y. J. Jeon, W. Y. Kim and H. S. Kim. 2005. HESAS: HERVs Expression and Structure Analysis System. *Bioinformatics.* **15**, 21, 1699-700.
- Lander, E.S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle and W. FitzHugh. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**, 860 - 921.
- Makalowski, W., G. A. Mitchell and D. Labuda. 1994. Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends Genet.* **10**, 188-193.
- Marin, D.B., M. C. Sewell and A. Schlechter. 2002. Alzheimer's disease. Accurate and early diagnosis in the primary care setting. *Geriatrics* **57**, 36-40.
- Mattson, M.P. 1997. Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiol. Rev.* **77**, 1081-1132.
- Nigumann, P., K. Redik, K. Matlik and M. Speck. 2002. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics* **79**, 628-634.
- Rajendran, L., M. Honscho, T. R. Zahn, P. Keller, K. D. Geiger, P. Verkade and K. Simons. 2006. Alzheimer's disease beta-amyloid peptides are released in association with exosomes. *Proc. Natl. Acad. Sci. U. S. A.* **25**, 11172-11177.
- Rall, S.C.Jr., K. H. Weisgraber and R. W. Ahley. 1982. Human Apolipoprotein E. *J. Biol. Chem.* **257**, 4171-4178.
- Selkoe, D. J. 1996. Amyloid -protein and the genetics of Alzheimer's disease. *J. Biol. Chem.* **271**, 18295 - 18298.
- Simons, C., M. Pheasant, I. V. Makunin and J. S. Mattick. 2005. Transposon-free regions in mammalian genomes. *Genome Res.* **16**,164-172.
- Sin, H.S., J. W. Huh, D. S. Kim, D. W. Kang, D. S. Min, T. H. Kim, H. S. Ha, H. H. Kim, S. Y. Lee and H. S. Kim. 2006. Transcriptional control of HERV-H LTR element of GSDML gene in human tissues and cancer cells. *Arch. Virol. (in press)*.
- Smit, A.F. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* **9**, 657 - 663.
- Sorek, R., G. Ast and D. Graur. 2002. Alu-containing exons are alternatively spliced. *Genome Res.* **12**, 1060-1067.
- Stromberg, K., E. M. Hansson, H. Laudon, S. Bergstedt, J. Naslund, J. Lundkvist and U. Lendahl. 2005. gamma-Secretase complexes containing N- and C-terminal fragments of different presenilin origin retain normal gam-

- ma-secretase activity. *J. Neurochem.* **95**, 880-890.
26. Tanzi, R.E., J. F. Gusella, P. C. Watkins, G. A. P. Bruns, P. St George-Hyslop, M. L. Van Keuren, D. Patterson, S. Pagan, D. M. Kurnit and R. L. Neve. 1987. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science* **235**, 880-884.
27. Ye, S., Y. Huang, K. Mullendorff, L. Dong, G. Giedt, E. C. Meng, F. E. Cohen, I. D. Kuntz, K. H. Weisgraber and R. W. Mahley. 2005. Apolipoprotein (apo) E4 enhances amyloid beta peptide production in cultured neuronal cells: apoE structure as a potential therapeutic target. *Proc. Natl. Acad. Sci. U. S. A.* **20**, 18700-18705.
28. Zhang, Z., S. Schwartz, L. Wagner and W. Miller. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203 - 214.