

암 관련 isoform switch 유전자에 대한 기능분석 연구 A Function Study on Cancer related Isoform Switch in TCGA

유 제 운, 김 대 수
한국생명공학연구원

Jea-woon Ryu, Dae-soo Kim
Korea Research Institute of Bioscience &
Biotechnology

요약

유전자는 alternative splicing으로 환경에 맞는 전사체를 발현함으로써 생명체의 복잡성을 야기한다. 하지만 대부분의 유전자는 dominant한 isoform 하나만이 주로 발현되며 그 외는 극히 미약하게 발현을 한다. 정상 대비 암조직에서 dominant하게 발현되는 isoform은 암의 성장과 발달에 영향을 끼칠 것으로 보인다. 이에 정상과 비교하여 암 특이적인 isoform, 즉 isoform switch가 일어나는 유전자를 찾기 위해 TCGA의 RNA-seq 데이터를 다운로드하여 isoform switch가 일어나는 유전자를 뽑았다. 12 조직에 대한 2,925 유전자는 기능분석을 통해 암과 밀접하게 연관되어 있음을 알 수 있었다.

I. 서론

유전자는 하나의 isoform 또는 여러 개의 isoform으로 구성되어 시간, 장소 등 조건에 따라 다양한 isoform 형태로 기능을 한다. Alternative splicing(선택적 스플라이싱) 방법이 이를 가능하게 하는 방법 중 하나이다. 유전자 레벨에서의 분석이 주된 방법으로 이용되고 있는데, isoform 레벨의 발현 프로파일 정보가 유전자 레벨의 발현 프로파일보다 보다 암에 대한 많은 정보를 준다는 연구결과가 발표되고 있다[1]. 최근 여러 논문에서 특정 isoform이 암 등 질병과 밀접한 연관되어 있음이 알려져 있다[2]. 또한 상황(condition)에 따른 유전자의 전사체 차별적 사용(isoform switch)이 암 유형 전반에 걸쳐 연구가 진행되고 있다[3].

이에 정상과 비교하여 암에서 isoform switch가 일어나는 유전자를 찾고 기능분석(function study)을 하고자 한다. 이를 통해 암 특이적으로 발현하는 isoform이 암의 성장과 발달에 어떻게 영향을 끼치고 있는지를 알 수 있게 됨으로써 암에 대한 이해도를 높일 수 있을 것이다.

II. 재료 및 방법

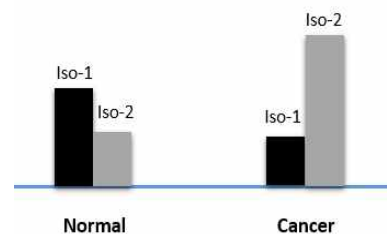
1. datasets

TCGA(The Cancer Genome Atlas)는 암에 관한 유전 변이, 발현, 메틸정보 데이터를 통합하여 생물정보 분석을 하고자 하는 목적으로 만들어진 대규모 암 유전체 연구 프로젝트이다. 각 조직별로 level 3의 isoform expression quantification RNA-seq 데이터를 TCGA

portal에서 다운로드 하였다. 조직 당 샘플 수가 10개 이상이며, 정상과 암 데이터 모두 가지고 있는 데이터를 이용하였다 (bladder, breast, colorectal, esophagus, head and neck, kidney, liver, lung, prostate, stomach, thyroid, uterus). 유전자 발현값은 TPM 값으로 변환하여 사용하였고, Isoform의 UCSC ID는 UCSC genome browser를 통해 유전자 symbol로 변환하였다.

2. Isoform switch

각 조직별로 여러 샘플이 가지는 발현값 중 median 값을 대표값으로 선정하여 정상조직 대비 암조직에서 차이나는 유전자를 선택하였다. Isoform switch를 보기위해서 하나 이상의 isoform을 갖는 유전자를 대상으로 진행하였고, 유의하게 차이나는 isoform을 찾기 위해 TPM 1이상인 isoform을 대상으로 진행하였다 (그림 1).



▶▶ 그림 1. Isoform switch

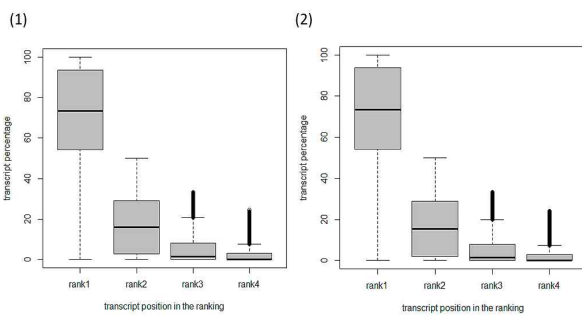
3. Function study

암에서 주로 나타나는 isoform에 대한 function study는 DAVID functional annotation을 통해 KEGG 경로 분석을 진행하였고, heat map은 R을 통해 진행하였다.

III. 본론

TCGA portal를 통해 다운받은 isoform expression quantification 데이터 중 12개의 조직에 대해서 분석을 진행하였다. 유전자 발현은 TPM 값으로 전환하여 사용하였다.

대부분의 단백질 코딩 유전자들은 하나의 dominant 전사체를 가지는데 정상뿐만 아니라 암에서도 major isoform이 전체 전사물에서 75% 정도의 발현 비중을 가지고 있음을 알 수 있었다 (그림 2). 이 결과로 유전자당 major/minor 한 isoform으로 크게 구분하여 분석하여도 무리가 없음을 알 수 있다.



▶▶ 그림 2. 정상(1)과 암조직(2)에서의 isoform percentage 분포 분석

정상 대비 암조직에서 유의하게 isoform switch가 일어나는 유전자를 찾았다 (표1). 이 리스트에는 단백질로 코딩되지 않는 psedogene들을 포함하고 있다. 각 조직별 isoform switch를 나타내는 유전자 리스트를 정리하였고 이를 모두 합한 전체 유전자 2,925개 중 약 60%는 두 개의 조직에서 중복되어 나타났고, 40%는 조직 특이적으로 나타나는 유전자임을 알 수 있었다 (data not shown). 모든 조직에서 나타나는 isoform의 경우 ADH1B, ANGPTL1, CNTFR, DPT, FOSB, FXYD1, RPL36A, UBE2C이 있다. 암 진단 마커로 알려져 있는 CNTF, DPT, FOSB와 암과 발현 혹은 polymorphism과 관련이 있다고 알려진 ADH1B, FXYD1, RPL36A, UBE2C, 그리고 암을 억제한다고 알려진 ANGPTL1이었다.

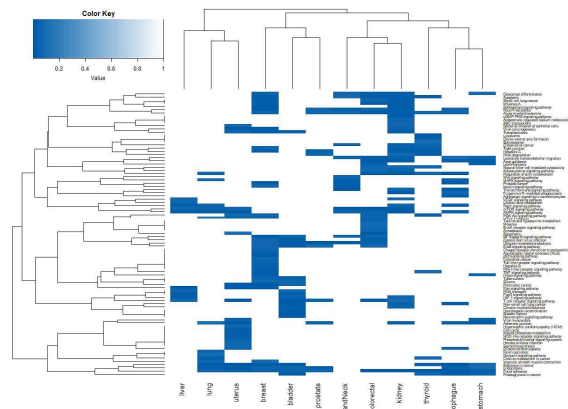
표 1. isoform switch

	Tissue	Isoform switch
1	Bladder	698
2	Breast	589
3	Colorectal	725
4	Esophagus	543
5	HeadandNeck	351
6	Kidney	753
7	Liver	307
8	Lung	539
9	Prostate	562
10	Stomach	472
11	Thyroid	663
12	Uterus	943

관여하는 KEGG 경로 분석을 진행한 결과 (그림 3) 조직별 특이하게 나타나는 경로를 확인할 수 있었다.

bladder cancer를 예를 들면 bladder cancer, homologous recombination, neurotrophin signaling pathway로 이미 bladder와 밀접하게 연관되어 있다고 알려진 경로이다.

공통적으로 대다수 조직(≥ 7 조직)에서 나타나는 경로는 focal adhesion, endocytosis, pathways in cancer, acute myeloid leukemia, AMPK signaling pathway이다. 암의 성장과 이동에 관여하고 세포내 에너지 항상성 유지에 센서 역할을 효소 및 signal transduction 등 전반적으로 암이 발달하는데 isoform switch가 일어남을 알 수 있었다.



▶▶ 그림 3. 암별 KEGG 경로 clustering 분석

IV. 고찰

정상 대비 암에서 특이하게 나타나는 isoform switch에 관여하는 유전자들에 대한 기능분석을 진행한 결과 암과 관련된 경로임을 확인할 수 있었다. 추후 연구로 DEGs (differential expressed genes)와의 비교를 통해 암에서의 isoform에 대한 심도 깊은 연구를 진행할 계획이다.

■ 참고 문헌 ■

- [1] Zhang ZF., Pal S., Bi Y., Tchou J., and Davuluri RV., "Isoform level expression profiles provide better cancer signatures than gene level expression profiles," *Genome Med.*, Vol 5, No. 4, pp. 33, 2013.
- [2] Oltean S., Bates DO., "Hallmarks of alternative splicing in cancer," *Oncogene.*, Vol 33, No. 46, pp. 5311-5318, 2014.
- [3] Vitting-Seerup K., Sandelin A., "The landscape of isoform switches in human cancers," *Mol Cancer Res.*, Vol. 15, No. 9, pp. 1206-1220, 2017.