creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

의학박사 학위논문

갑상선 유두 암 환자의 나이에 따른 예후와 관련된 유전체 발현 양상 분석: cBioPortal을 사용한 Pan-Cancer Atlas 분석

Analysis of Gene Expression Signatures Identifying Age-Associated Mortality in Papillary Thyroid Carcinoma Patients: an Analysis of Pan-Cancer Atlas using cBioPortal

울 산 대 학 교 대 학 원

의 학 과

오 혜 선

갑상선 유두 암 환자의 나이에 따른 예후와 관련된 유전체 발현 양상 분석: cBioPortal을 사용한 Pan-Cancer Atlas 분석

지 도 교 수　　김 태 용

이 논문을 의학박사 학위 논문으로 제출함

2022년 2월

울 산 대 학 교 대 학 원

의　　학　　과

오 혜 선

오혜선의 의학박사학위 논문을 인준함

심사위원　　백정환　(인)

심사위원　　김태용　(인)

심사위원　　이우창　(인)

심사위원　　김태혁　(인)

심사위원　　전민지　(인)

울 산 대 학 교 대 학 원

2022년 2월

국 문 요 약

배경: 나이는 갑상선 암의 생존에 있어 가장 중요한 예후 인자 중 하나로, 여러 갑상선 암 위험도 평가 체계에서 나이가 주요한 요소로 고려되어 왔다. 하지만 이에 영향을 주는 유전자에 대한 연구 결과는 부족한 실정이다.

연구방법: 본 연구는 cBioPortal를 이용하여 2018년도 새롭게 발표된 Pan-Cancer Atlas 데이터를 활용하여 20세 이상 갑상선 유두 암 환자(490명)의 임상 및 유전자 정보를 얻은 뒤, 55세 미만(325명)과 55세 이상(165명)의 두 군으로 환자들을 구분하였다. 두군 간에 유의하게 발현량이 차이가 나는 mRNA (812개)를 추출 하고 R을 활용하여 812개 모든 유전자들에 대하여 55세를 기준으로 무진행 생존(PFS)과 전체생존기간(OS)의 차이를 계산 하였고, 그 중 p value가 통계적으로 유의하게 차이가 나는 유전자만 추출 하였다. 이후, 각 유전자에 대하여 갑상선 분화도(thyroid differentiation score (TDS), *BRAF*^{V600E}*-RAS* score (BRS), extracellular signal-regulated kinase 점수(ERK 점수))와 상관관계를 분석 하였다.

결과: 환자들의 평균 나이는 47.9세 였고 그 중 165명(33.7%)가 55세 이상이었다. *BRAF* 돌연변이는 254명(51.8%)에서, *RAS* 돌연변이는 48(9.8%)에서 발견되었다. 55세 이상 환자가 55세 미만 환자와 비교 시 통계적으로 유의하게 생존 위험도가 증가되어 있었다(hazard ratio = 35.1, $p < 0.001$). 총 14개의 유전자(*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*)가 55세 이상의 병의 재발 및 생존 위험도 증가와 관련된 유전자로 확인되었다. 각 유전자와 갑상선 분화도와 상관관계 분석을 하였다; TDS와 양의 상관관계를 보이는 유전자(*F12, NRSN2, PPP2R1B, RRP36*), 음의 상관관계를 보이는 유전자(*KBTBD2*), BRS와 양의 상관관계를 보이는 유전자(*ABCF1, FLAD1, IARS2, RRP36*), 음의 상관관계를 보이는 유전자(*ENTPD1, HCG11,*

KBTBD2, QSOX2, RINL, ZFAND5), ERK 점수와 양의 상관관계를 보이는 유전자 (*ENTPD1, HCG11, IARS2, KBTBD2, RINL, ZFAND5*), 음의 상관관계를 보이는 유전자 (*ABCF1, NRSN2*).

결론: 본 연구에서는 cBioPortal를 이용하여 Pan-Cancer Atlas 데이터를 통해 유두 갑상선 암 환자에서 나이와 관련된 예후에 영향을 미치는 유전자를 확인하였다. 또한 이들과 갑상선 분화도 상관성을 입증하였다. 향후 해당 유전자들을 임상 치료 및 예후 평가에 적용할 수 있는 후속 연구가 필요하겠다.

중심단어: 갑상선 유두 암, Pan-Cancer Atlas, cBioPortal, 나이, 유전자 분석

차 례

국문요약 · i

차례 · iii

표, 그림, 약어 목록 · v

서론 · 1

1. 나이에 따른 갑상선 암의 예후 · · · · · · · · · · · · · · · · · 1

2. 갑상선 유두 암의 유전자 변이 · · · · · · · · · · · · · · · · · 2

3. 암 유전체 아틀라스와 Pan-Cancer Atlas · · · · · · · · · · · · 4

4. cBioPortal · 7

5. 연구 목적 · 8

연구대상 및 연구방법 · 9

I. cBioPortal을 활용한 Pan-Cancer Atlas 데이터 수집 · · · · · 9

1. 환자들의 임상 정보 수집 · 11

2. 갑상선 분화도 계산 (TDS, BRS, ERK 점수) · · · · · · · · · · 12

1) TDS (Thyroid Differentiation score) · · · · · · · · · · · 12

2) BRS (B*RAF*^{V600E}-*RAS* score) · · · · · · · · · · · · · · · · · 14

3) ERK (extracellular signal-regulated kinase) 점수 · · · · 16

3. 나이에 따른 예후와 관련된 유전체 분석 · · · · · · · · · · 18

4. 추출된 유전자와 갑상선 분화도와 상관관계 분석 · · · · · · 19

II. 통계학적 분석 · 20

연구결과 · 21

1. 연구 환자들의 임상 및 병리학적 특성 · · · · · · · · · · · · 21

2. 나이에 따른 예후와 관련된 유전자 · · · · · · · · · · · · · 23

3. 추출된 유전자와 갑상선 분화도와 상관관계 분석 · · · · · · · 24

고찰 · 30

결론 · 34

참고문헌 · 35

영문요약 · 39

보충자료: 14개 유전자의 간단한 설명 · · · · · · · · · · · · · · 41

표, 그림, 약어 목차

그림1. 정상 갑상선과 갑상선 유두 암 사이 MAPK 신호전달 체계 차이 비교 2

그림 2. Pan-Cancer Atlas의 종양유전학적 개요 · · · · · · · · · · · · · 4

그림 3. 암 유전체 아틀라스 데이터와 Pan-Cancer Atlas 데이터의 mRNA-Seq 처리 비교 · 5

그림 4. 데이터 분석 알고리즘 · 10

표 1. TDS (Thyroid Differentiation Score)에 사용된 16개의 유전자 · 13

그림 6. BRS (*BRAF*^{V600E}-*RAS* score)에 사용된 71개의 유전자 · · · · · · · 15

그림 7. ERK (extracellular signal-regulated kinase) 점수에 사용된 52개의 유전자 · 17

그림 8. 나이에 따른 예후와 관련된 유전체 분석 간략 모식도 · · · · · 18

표 2. 유두 갑상선 암 환자들의 임상 및 병리학적 특성 · · · · · · · · 22

그림 9. 나이에 따른 생존 곡선 · 23

표 3. 추출된 유전자와 갑상선 분화도와 상관관계 분석 표 · · · · · · · 25

그림 10. Thyroid differentiation score와 유의한 상관 관계를 보이는 유전자 분석 · 26

그림 11. *BRAF*^{V600E}-*RAS* score 와 유의하게 양의 상관 관계를 보이는 유전자 분석 · 27

그림 12. *BRAF*^{V600E}*-RAS* score 와 유의하게 음의 상관 관계를 보이는 유전자 분석 · 28

그림 13. Extracellular signal-regulated kinase score와 유의한 상관 관계를 보이는 유전자 분석 · 29

서론

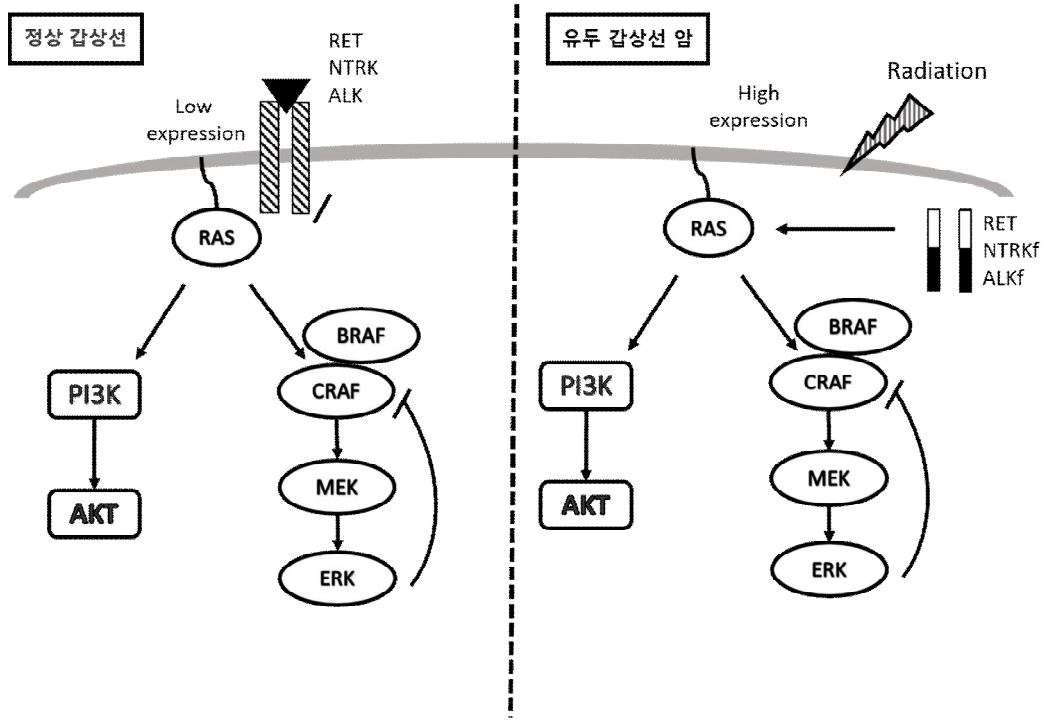
1. 나이에 따른 갑상선 암의 예후

나이는 갑상선 암의 생존에 있어 가장 중요한 예후 인자 중 하나로, 여러 갑상선 암 위험도 평가 체계에서 나이가 주요한 요소로 고려되어 왔다[1-3]. 미국 암 연합회(AJCC, American Joint Committee on Cancer)는 지난 2016년도에 8판 TNM (Tumor, node, metastasis) 병기 체계를 발표하였다. 7판과 8판의 가장 큰 변화 중 하나가 바로 병기 설정의 경계 나이를 45세에서 55세로 상향 조정 한 것이다[4-6]. 갑상선 암 환자의 나이가 증가할수록 예후가 나빠진다는 것은 여러 연구를 통해 알려져 있는 사실로[7-9], 과거 연구에서는 45세에서 55세로 나이를 상향 조정 했을 때 환자의 사망률 예측이 더 계층화 된다는 것이 보고된 바 있다[5,10]. 또한 Adam등은 미국의 SEER (Surveillance, Epidemiology, and End Results) 자료를 바탕으로 갑상선 암 환자의 나이가 증가함에 따라 사망률이 선형 관계로 증가된다고 보고하였다[9].

2. 갑상선 유두 암의 유전자 변이

갑상선 유두 암(PTC, papillary thyroid carcinoma)의 치료를 위해 다양한 유전자 연구가 이루어 져 왔으며 과거 연구들을 통해 *BRAF*^{V600E}와 *RAS*유전자의 점 돌연변이와 *neutrophic receptor tyrosine kinase1 (NTRK1)*와 *RET* 유전자 융합을 포함하는 *mitogen-activated protein kinase (MAPK)* 신호전달 체계와 관련된 유전자들의 체세포 변이가 분화 갑상선 암에서 다 빈도로 관찰된다는 사실은 잘 알려져 있다(그림 1)[11-16].

그림 1. 정상 갑상선과 갑상선 유두 암 사이 MAPK 신호전달 체계 차이 비교[16]



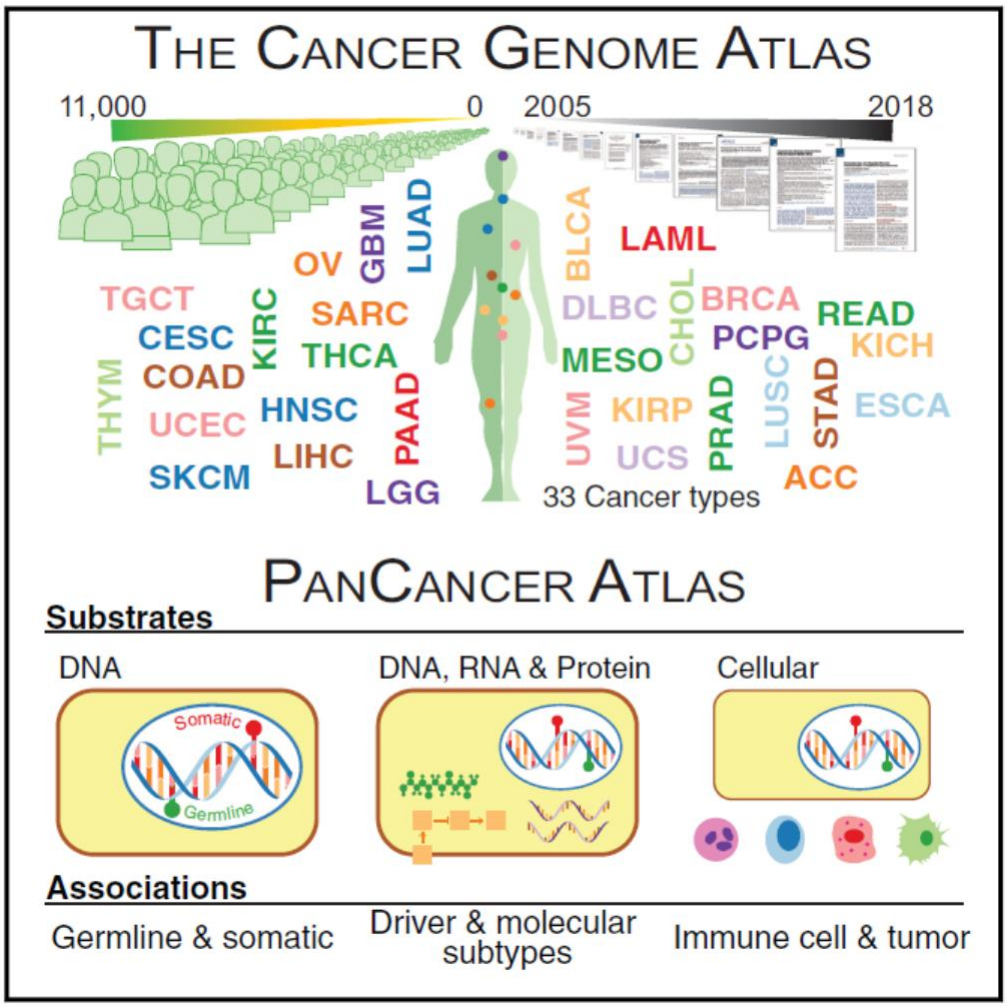
정상 갑상선에서는 *RET, neutrophic receptor tyrosine kinase (NTRK)*와 *ALK* 수용체 발현이 매우 적어 mitogen-activated protein kinase (MAPK)와 PI3K의 하부경로 활성화를 위해 작용제가 필요하다(좌측). 방사선등에 노출시 세포질 내에 키메릭 수용체가 비정상적으로 존재하게 되고 수용체 작용제와 상관없이 신호전달체계를 활성화 시킨다 (우측). 이러한 유전적 변이는 갑상선 유두 암에서 관찰된다. →는 활성화를 ⊣는 억제를 나타낸다.

2014년도에 암 유전체 아틀라스(TCGA, The Cancer Genome Atlas Research)를 이용한 갑상선 유두 암의 통합된 게놈 특성화 연구 결과가 발표 되었고 이 연구를 통해 갑상선 유두 암의 알려지지 않은 유전자 발암 인자가 25%에서 3.5%로 줄어들게 되었다[17]. 갑상선 유두 암의 *BRAF*^{V600E}와 *RAS* 변이 유전자들은 서로 상호 배타적으로 작용하는데, *BRAF*^{V600E} 변이에 의해 발생한 암은 extracellular signal-regulated kinase (ERK)와 RAF의 음성 피드백에 영향을 받지 않아 MAPK 신호체계를 과 발현 시키고, 반대로 *RAS* 변이와 *RTK* 융합으로 인해 발생한 종양은 ERK 피드백에 영향을 받아 MAPK 신호 체계의 발현을 줄이게 된다[18,19]. 그리고 이러한 차이가 결국 종양의 표현형에 영향을 주게 된다. 이에 암 유전체 아틀라스 연구에서는 각 종양을 *BRAF*^{V600E} like 갑상선 유두 암과 *RAS* like 갑상선 유두 암으로 구분하여 이 관계를 밝히고자 하였다. 해당 연구에서는 이를 확인하기 위해 BRS (*BRAF*^{V600E}-*RAS* score)를 이용하여 *BRAF*^{V600E} like과 *RAS* like 종양을 구분하였고, 갑상선 분화도와 관계를 분석하기 위해 갑상선 대사와 기능과 관련된 16개 유전자의 발현량을 이용하여 계산한 TDS (Thyroid differentiation score)라는 개념을 도입하였다. 이를 통해 TDS와 BRS가 양의 상관관계를 가진다는 것을 확인하였다(Spearman = 0.78, $p < 0.001$)[19]. 또한 *BRAF*^{V600E}와 *RAS* 유전 변이의 하부단계 경로에 주는 영향을 분석하기 위해 MAPK 신호전달체계의 발현 정도를 평가하는 52개의 유전자 발현량을 이용하여 ERK 점수를 계산하였고 *BRAF*^{V600E} like 갑상선 유두 암은 해당 신호전달체계가 과발현 되어 있다는 것을 확인하였다[18,19].

3. 암 유전체 아틀라스와 Pan-Cancer Atlas

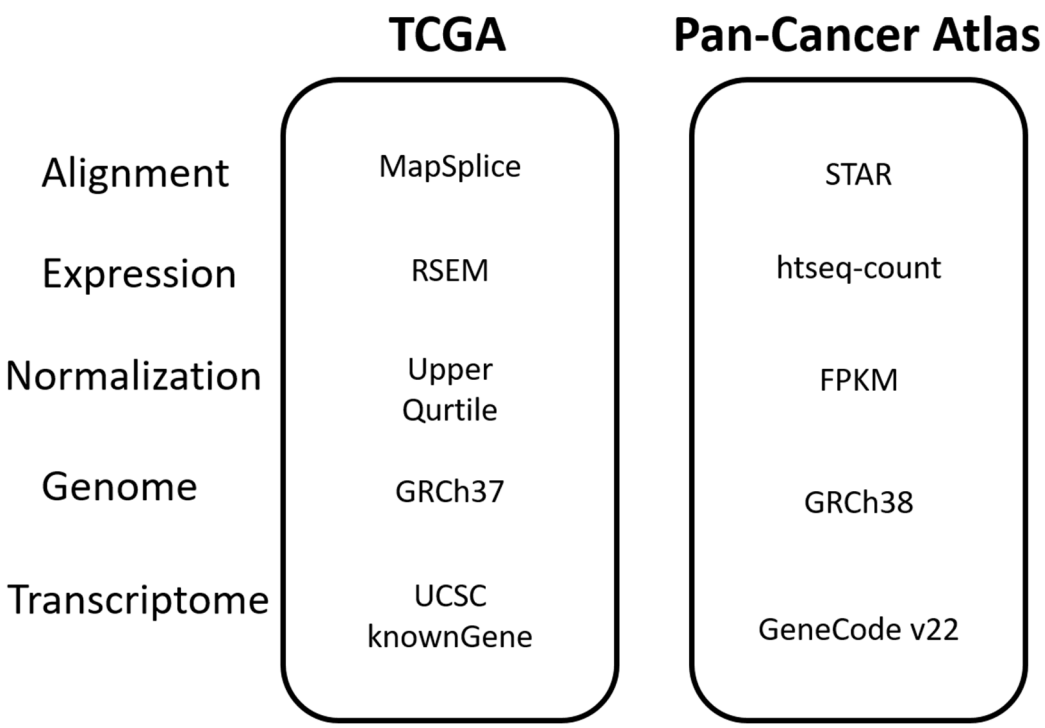
2018년도에 국제 암 유전체 컨소시엄(ICGC, International Cancer Genome Consortium)과 암 유전체 아틀라스에서 Pan-Cancer Atlas데이터를 발표하였다[20]. Pan-Cancer Atlas에서는 유전자 융합(fusion), 스플라이싱(splicing), 홀배수체(aneuploidy)등의 체세포 돌연변이와 영상분석 그리고 면역시스템 종양유전학 정보가 포함되었다(그림 2)[21].

그림 2. Pan-Cancer Atlas의 종양유전학적 개요[21]



암 유전체 아틀라스 데이터와 Pan-Cancer Atlas 데이터의 메신저 RNA (mRNA) 발현 정보 차이를 비교해 보면 크게는 염기 서열의 배열, 정량화, 정상화 그리고 mRNA-Seq의 참조 데이터에 차이가 있다(그림 3)[22].

그림 3. 암 유전체 아틀라스 데이터와 Pan-Cancer Atlas 데이터의 mRNA-Seq 처리 비교[22]



암 유전체 아틀라스 데이터와 Pan-Cancer Atlas 데이터 사이의 생체정보학 비교개요

염기 서열을 분석하는데 있어 더 발전된 기구(Illumina Gnome Analyzer GA에서 HiSeq로 변화하였다)와 화학 키트가 사용되었다. 또한 mRNA-Seq 데이터는 기존 암 유전체 아틀라스 데이터는 hg19 참고 게놈 데이터를 사용한데 반해, Pan-Cancer Atlas에서는 hg38 참고 게놈 데이터를 사용하여 분석하였다. 기존 암 유전체 아틀라스 데이터는 발현값의 정상화를 위해 사분위정상화 방법을 사용한데 반해, Pan-Cancer Atlas에서는 FPKM (Fragments Per Kilobase of exon per Million) 예측 값을 활용하였다.

4. cBioPortal

cBioPortal[23] (http://cbioportal.org)은 암의 게놈 정보를 공개적으로 이용 가능한 자료로 제공하며, 현재 5000개가 넘는 종양 샘플을 20개의 종양 연구(5개의 연구 데이터와 15개의 TCGA 데이터)로부터 데이터를 얻어 제공하고 있다. 이는 Memorial Sloan-Kettering Cancer Center (MSKCC)에서 개발되었고, cBioPortal의 웹 서비스는 MSKCC의 Center for Molecular Oncology에서 GitHub의 허가를 통해 공개 정보로 관리하고있다. 각 데이터들은 유전자 레벨에서 저장된 뒤 무진행 생존(PFS, progression free survival)과 전체생존기간(OS, overall survival)등의 임상 정보가 통합되어 저장되어 있다. 이 cBioPortal을 통해 복잡한 유전자 정보와 암 연구자들 사이의 장벽이 낮아 졌고 빠르고 높은 수준의 연구가 가능하게 되었다.

5. 연구 목적

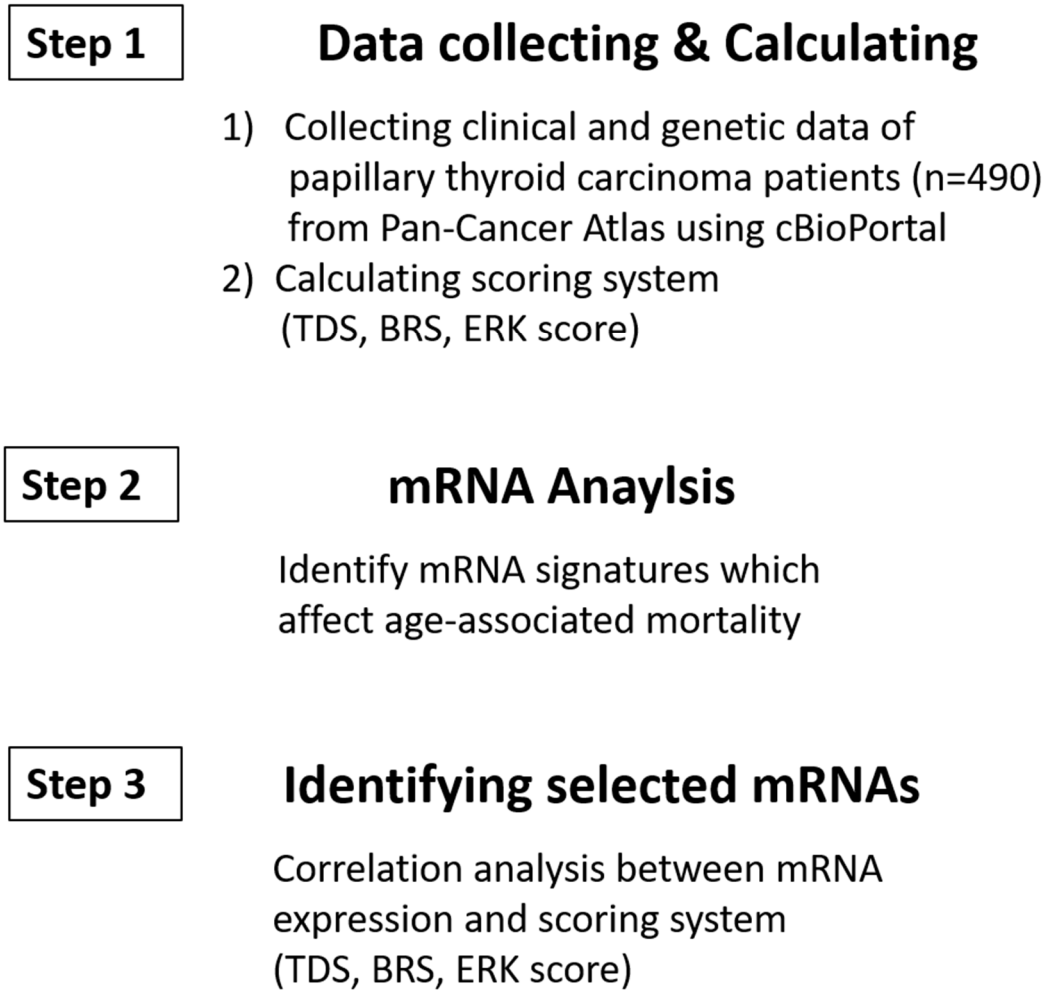
BRAF^{V600E} 돌연변이가와 *TERT* 돌연변이가 고령에서 높은 빈도로 발견되는 것은 여러 연구 결과를 통해 알려져 있다[24,25]. 또한 Min 등은 *BRAF*^{V600E} 돌연변이가 있는 환자에서 나이와 관련된 사망률이 선형 관계를 이루지만 *BRAF*^{V600E} 가 없는 환자들에서는 확인되지 않는다는 것을 보고한 바 있다[25]. 하지만 *BRAF*^{V600E} 와 *TERT* 돌연변이 이외에 나이와 갑상선 암의 나쁜 예후에 영향을 미치는 유전자에 대한 연구 결과는 적다[7-9,24,25]. 이에 본 연구에서는 cBioPortal (https://www.cbioportal.org/datasets)을 이용하여 Pan-Cancer Atlas[20] 데이터를 활용해 갑상선 유두 암 환자에서 나이와 관련된 예후에 영향을 미치는 유전자를 분석하고자 한다.

연구대상 및 연구방법

I. cBioPortal을 활용한 Pan-Cancer Atlas 데이터 수집

본 연구는 cBioPortal (https://www.cbioportal.org/datasets)을 이용하여 국제 암 유전체 컨소시엄과 암 유전체 아틀라스에서 2018년도 발표한 Pan-Cancer Atlas[20] 데이터를 활용하여 시행한 연구로, 1. 환자들의 임상 정보 수집, 2. 갑상선 분화도 계산 (TDS, BRS, ERK 점수), 3. 나이에 따른 예후와 관련된 유전체 분석, 4. 추출된 유전자와 갑상선 분화도와 상관관계 분석의 과정으로 이루어졌다(그림 4).

그림 4. 데이터 분석 알고리즘



TDS, thyroid differentiation score; BRS, *BRAF*^{V600E}*-RAS* score; ERK, extracellular signal-regulated kinase

1. 환자들의 임상 정보 수집

R studio와 R (version 3.5.1, R Foundation for Statistical Computing, Vienna, Austria; http://www..r-project.org/)에서 Bioconductor의 cBioPortal 패키지를 활용하여 “암 유전체 아틀라스, Pan-Cancer Atlas” 중 갑상선 암 환자들의 데이터를 API (Application Program Interface)를 통해서 접근하여 추출 하였다. 총 500명의 갑상선 암 환자 중 20세 미만의 10명 환자가 제외되어 490명의 환자가 분석에 활용되었다. 갑상선 암 환자들의 병기 설정에는 미국 암 연합회(AJCC) 4판(3명), 5판(26명), 6판(126명), 7판(345명)이 각각 사용되었다. 본 연구는 공개적으로 이용 가능한 자료를 이용하여 시행된 연구로 서울아산병원의 임상연구심의위원회의 심의 면제를 획득하였다.

2. 갑상선 분화도 계산 (TDS, BRS, ERK 점수)

1) TDS (Thyroid Differentiation score)

TDS는 2014년도에 암 유전체 지도 연구회(TCGA Research Network)에서 유두 갑상선 암 환자의 암 유전체 아틀라스를 분석한 논문에서[19] 처음으로 제시된 개념으로 16개의 갑상선 대사와 기능에 관련된 유전자들의 메신저 RNA (mRNA)의 발현량을 이용하여 계산하였다(표 1). 그 값이 높을수록 갑상선 분화도가 높은 것을 의미한다.

$$TDS = \text{Mean of } Log_2(\text{Fold Change}) \text{ across 16 genes}$$

표1. TDS (Thyroid Differentiation Score)에 사용된 16개의 유전자[19]

	Gene symbol	Gene full name	Brief function
1	*DIO1*	deiodinase, iodothyronine, type I	activates thyroid hormone by converting the prohormone thyroxine (T4) by outer ring deiodination (ORD) to bioactive 3,3'5'-triiodothyronine (T3)
2	*DIO2*	deiodinase, iodothyronine, type II	activates thyroid hormone by converting the prohormone thyroxine (T4) by outer ring deiodination (ORD) to bioactive 3,3'5'-triiodothyronine (T3)
3	*DUOX1*	dual oxidase 1	involved in synthese of thyroid hormone
4	*DUOX2*	dual oxidase 2	involved in synthese of thyroid hormone
5	*FOXE1*	forkhead box E1	thyroid transcription factor which likely plays a crucial role in thyroid morphogenesis
6	*GLIS3*	GLIS family zinc finger 3	functions as both a repressor and activator of transcription and is specifically involved in the development of pancreatic beta cells, the thyroid, eye, liver and kidney
7	*NIKX2-1*	NK2 homeobox1	thyroid-specific transcription factor that binds to the thyroglobulin promoter and regulates the expression of thyroid-specific genes
8	*PAX8*	paired box 8	transcription factor involved in thyroid follicular cell development and expression of thyroid specific genes
9	*SLC26A4*	solute carrier family 26 (anion exchanger), member 4	iodide transmembrane transporter activity
10	*SLC5A5*	solute carrier family 5 (sodium/iodide cotransporter), member 5	sodium:ioide symporter activity, responsible for the uptake of iodine in the thyroid
11	*SLC5A8*	solute carrier family 5 (sodium/monocarboxylate cotransporter), member 8	transport iodide by a passive mechanism
12	*TG*	thyroglobulin	Substrate for the synthesis of thyroxine and triiothyronine as well as the storage of the inactive forms of thyroid hormone and iodine
13	*THRA*	thyroid hormone receptor, alpha	nuclear hormone receptor for triiodothyronine
14	*THRB*	thyroid hormone receptor, beta	nuclear hormone receptor for triiodothyronine
15	*TPO*	thyroid peroxidase	iodination of tyrosine residues in thyroglobulin and phenoxy-ester formation between pairs of iodinated tyrosines to generate the thyroid hormone, thyroxine and triiodothyronine
16	*TSHR*	thyroid stimulating hormone receptor	receptor for thyrothropin and a major controller of thyroid cell metabolism

2) BRS (*BRAF*^{V600E}-*RAS* score)

유두 갑상선 암에서 가장 흔히 발견되는 유전자 돌연변이인 *BRAF*^{V600E}와 *RAS* 돌연변이를[26] 이용하여 계산한 점수이다. 암 유전체 지도 연구회에서 유두 갑상선 암 환자의 암 유전체 아틀라스를 분석한 논문에서 제시한 71개의 유전자를 활용하였다(그림 6)[19]. *BRAF*^{V600E} 돌연변이가 있는 종양을 B군으로, *RAS* 돌연변이가 있는 종양을 R군 나누었다. 이후 각 군의 71개 mRNA 값의 중심값을 각각 c(B), c(R)로 계산한 뒤 c(B)와 c(R)에서 각 종양의 벡터 거리의 차이를 이용하여 계산하였다. BRS 값은 +이면 *RAS* like 이며 -는 *BRAF*^{V600E} like 종양으로 해석 할 수 있다.

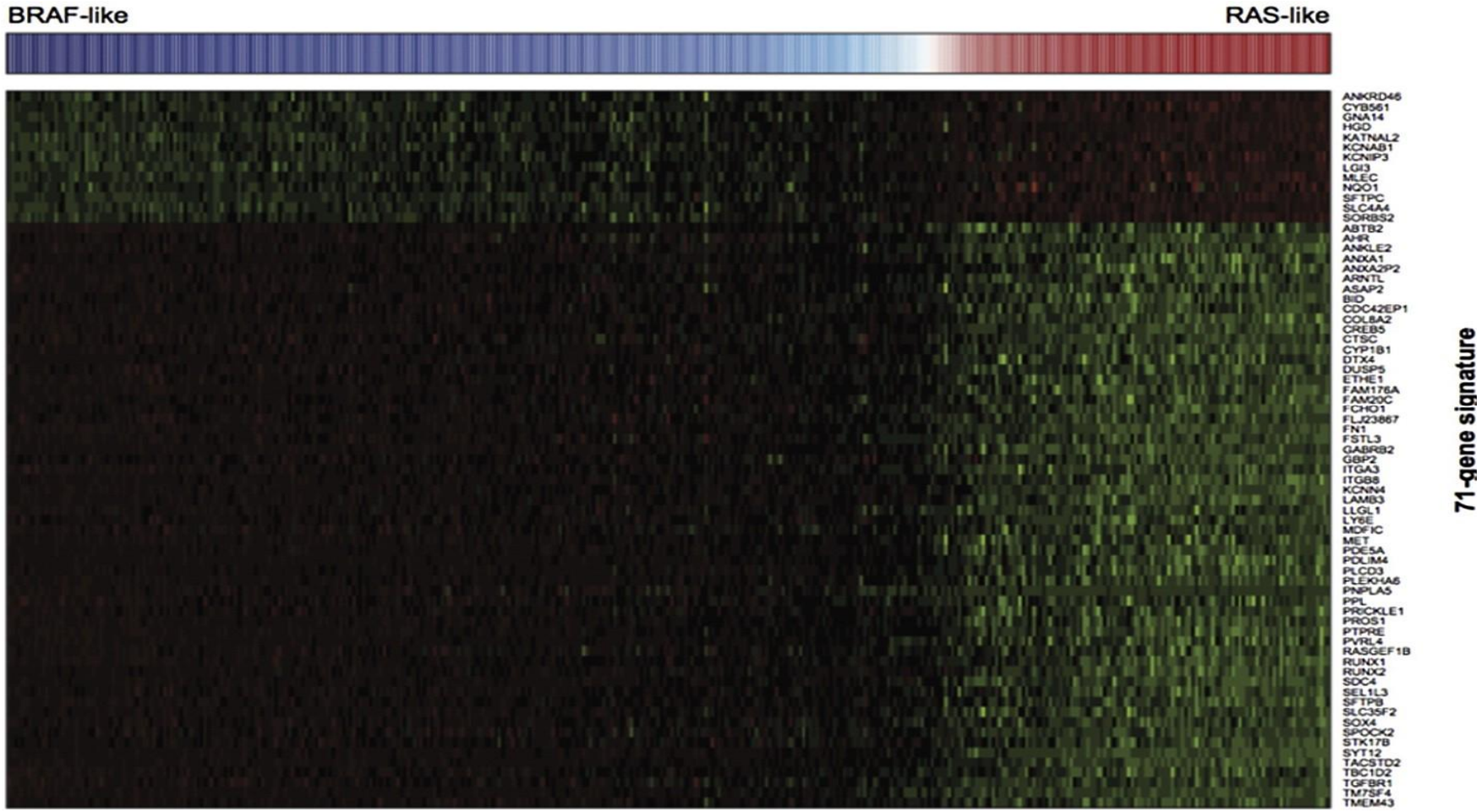
$$BRS(t) = |v(t)-c(B)|_2 - |v(t)-c(R)|_{-2}$$

* v(t): 종양(t)의 71개 유전자의 mRNA 발현 값 벡터

* $|x-y|_2$: 각 벡터 값을 유클리드 거리를 이용하여 정상화하여 계산

*BRS(t)를 계산한 뒤 최소-최대값 정상화를 이용하여 정상화함

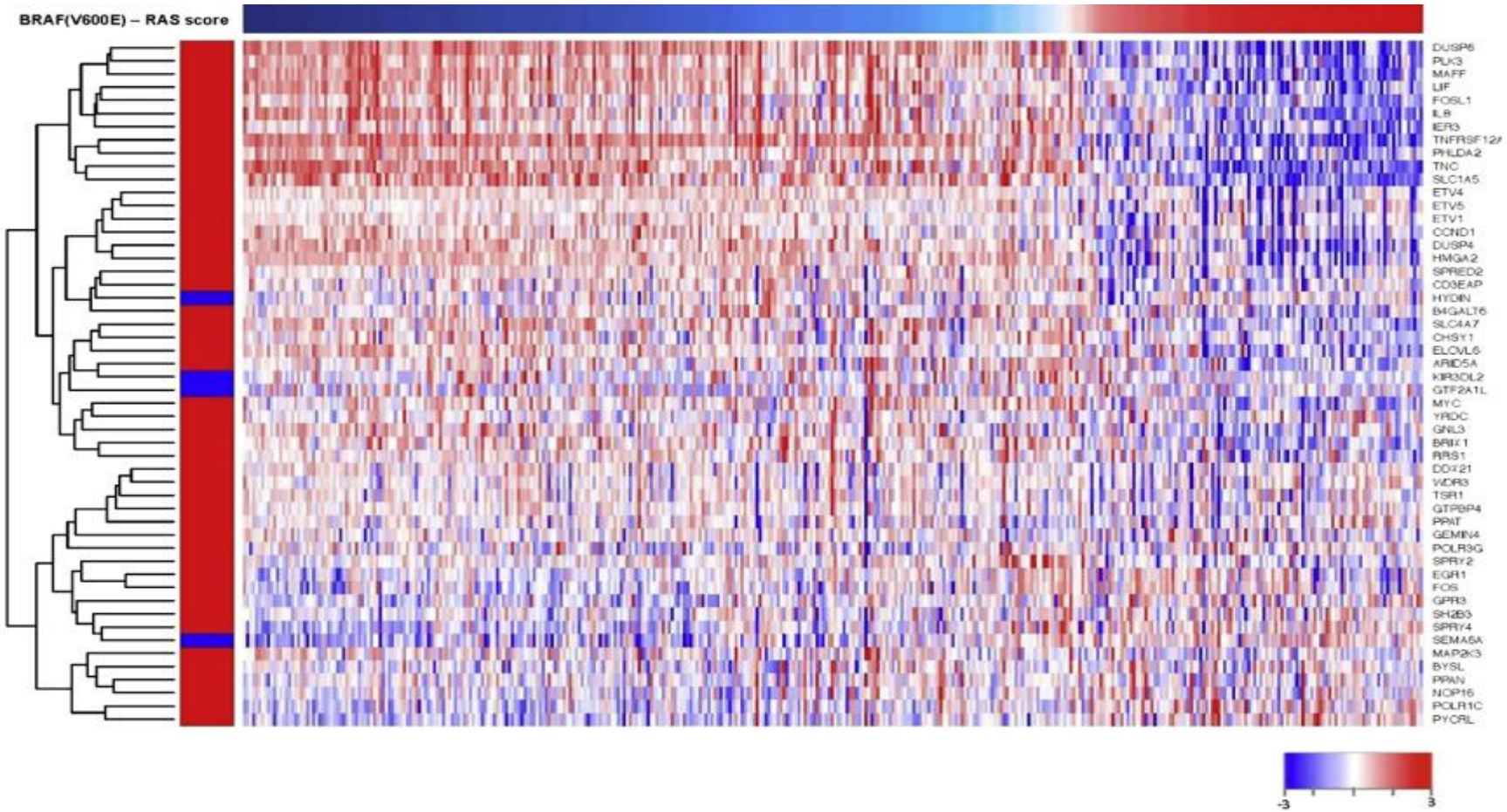
그림 6. BRS (*BRAF*^{V600E}-*RAS* score)에 사용된 71개의 유전자[19]



3) ERK (extracellular signal-regulated kinase) 점수

ERK 점수는 Pratilas등이 2009년도에 보고한 MEK (mitogen-activated protein kinase) 억제 치료에 반응이 있는 것으로 알려진 52개의 mRNA 유전자 발현을 이용하여 계산하였다(그림 7)[18,26]. 52개 mRNA 유전자의 Z 점수 값을 cBioPortal에서 추출 하여 MEK 억제 치료로 인해 발현이 억제되는 유전자를 A 그룹(48개), 발현이 증가되는 유전자를 B 그룹(4개)으로 구분하였다[19]. A 그룹의 Z 점수 값의 합에서 B 그룹의 Z 점수 값을 뺀 값으로 ERK 점수를 계산하였다.

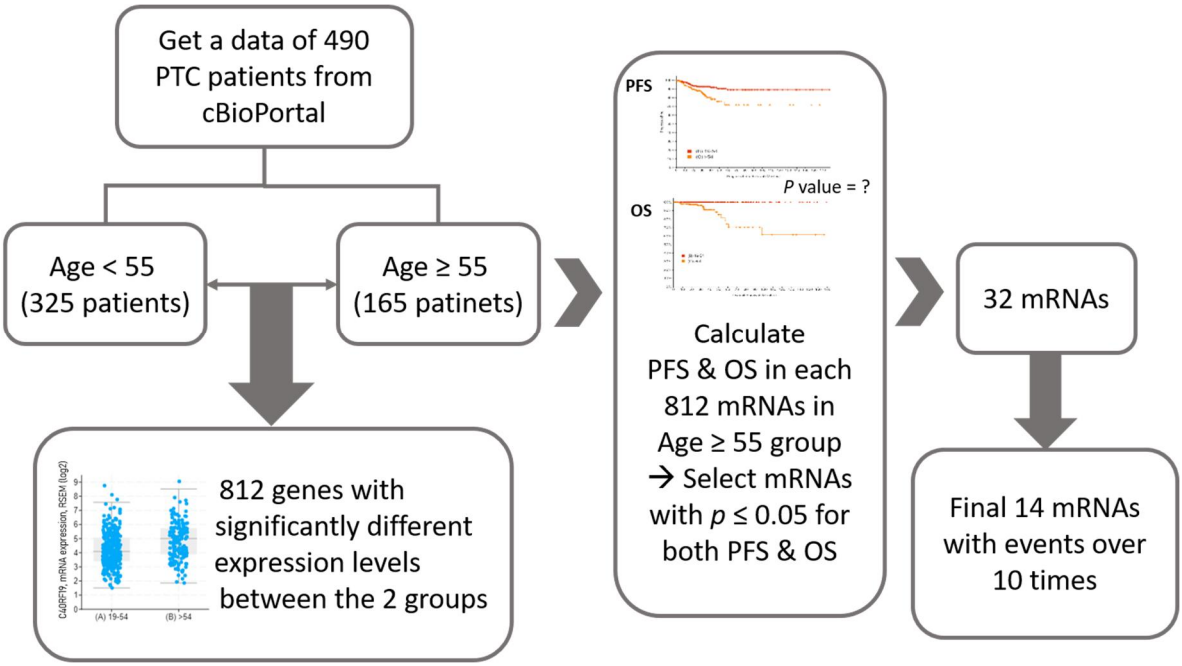
그림 7. ERK (extracellular signal-regulated kinase) 점수에 사용된 52개의 유전자[19]



3. 나이에 따른 예후와 관련된 유전체 분석

cBioPortal (https://www.cbioportal.org/datasets)에서 갑상선 암(암 유전체 아틀라스, Pan-Cancer Atlas)를 선택하여 데이터를 수집 하였다. 총 500명의 환자 중 20세 미만의 10명 환자를 제외한 뒤, 55세 미만(325명) 과 55세 이상(165명)의 두 군으로 환자들을 구분한 뒤, 두군 간에 유의하게 발현량이 차이가 나는 mRNA (812개)를 추출 하였다. 이후 R을 활용하여 812개 모든 유전자들에 대하여 55세 이상 환자에서 무진행 생존(PFS)과 전체생존기간(OS)의 차이를 계산하였고, 그 중 *p* value가 통계적으로 유의하게 차이가 나는 유전자만 추출 하였다. 총 32개의 유전자가 확인 되었고 그 중 병의 진행 혹은 사망의 이벤트가 10회 이하인 18개 유전자는 분석에서 제외하고 총 14개의 유전자가 분석에 활용되었다(그림 8).

그림 8. 나이에 따른 예후와 관련된 유전체 분석 간략 모식도



PFS, progression free survival; OS, overall survival

4. 추출된 유전자와 갑상선 분화도와 상관관계 분석

R을 활용하여 15개 유전자와 갑상선 분화도(TDS, BRS, ERK 점수) 사이의 상관관계를 분석하였다.

II. 통계학적 분석

모든 통계 분석은 R (version 3.5.1, R Foundation for Statistical Computing, Vienna, Austria; http://www..r-project.org/)과 R library package와 cBioPortalData (https://bioconductor.org/packages/release/bioc/html/cBioPortalData.html) 패키지를 사용하여 분석하였다. 연속변수는 중앙값과 사분위수 범위(IQR, interquartile range)로 표현되었고, 범주형 변수는 해당하는 수와 백분율로 표현되었다. 추출된 유전자들과 갑상선 분화도(TDS, BRS, ERK 점수) 사이의 상관관계 분석에는 Spearman 상관 분석이 사용되었다. 생존 분석에는 Kaplan-Meier 곡선과 유의성 확인을 위해서 log-rank 검정이 사용되었다. 모든 p 값은 양쪽 검정 하고, 0.05 미만의 p value 가 통계적 유의성을 가진다 판단하였다.

연구결과

1. 연구 환자들의 임상 및 병리학적 특성

총 490명의 유두 갑상선 암 환자들의 임상병리학적 특성은 표 2에 나와 있다. 전체 환자들의 평균 나이는 47.9세이며 165명(33.7%)이 55세 이상이다. 348명(71.0%) 환자가 순수 유두 갑상선 암 환자이며, 164명(33.5%) 환자가 TNM 병기 3-4기로 확인되었다. 218명(44.5%) 환자에서 림프절 전이가 확인되었고 림프절 전이 정보가 명확하지 않은 환자가 총 49명(10.0%)이었다. 8명(1.6%) 환자가 원격 전이가 있었으며 254명(51.8%)의 환자가 *BRAF* 돌연변이가 있었고 48명(9.8%)명의 환자들이 *RAS* 돌연변이가 확인되었다.

표 2. 유두 갑상선 암 환자들의 임상 및 병리학적 특성

	Total *n* = 490
Age at diagnosis (yrs)	47.9 ± 15.3
≥ 55	165 (33.7)
Female sex	358 (73.1%)
Histologic type of PTC	
Classical	348 (71.0%)
Follicular	99 (20.2%)
Tall cell	36 (7.3%)
Other	7 (1.4%)
TNM stage	
I/II	326 (66.5%)
III/IV	164 (33.5 %)
N stage	
N0	223 (45.5%)
N1	218 (44.5%)
Nx	49 (10.0%)
Distant metastasis	8 (1.6%)
BRAF mutation*	254 (51.8%)
RAS mutation**	48 (9.8%)

PTC, papillary thyroid carcinoma; T, tumor; N, node; M, metastasis.

Continuous variables are presented as means (standard deviation).

Categorical variables are presented as numbers (percentages).

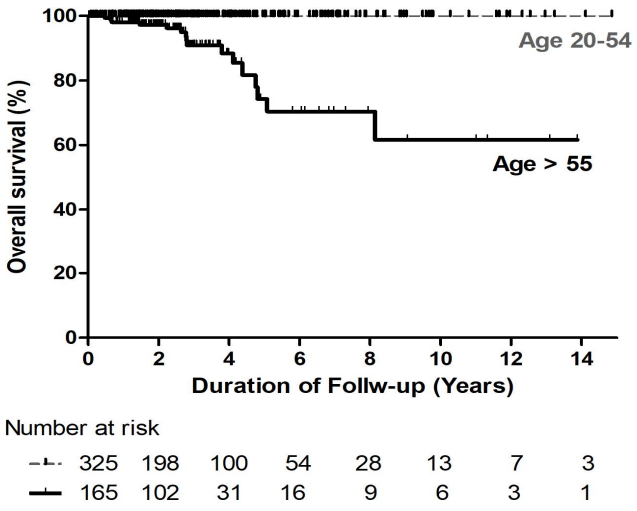
**BRAF* mutation includes $BRAF^{V600E}$, ^{K601E}, indels, fusions, focal deletions

** *RAS* mutation includes *NRAS, HRAS, KRAS*

2. 나이에 따른 예후와 관련된 유전자

중간 값 31.0(IQR 17.6 – 49.0)개월의 추적 관찰 기간 동안, 총 45(9.1%)명의 환자가 갑상선 암이 재발 또는 진행하였고(26(5.3%)명은 55세이상 환자) 16명의 환자(모두가 55세 이상)가 갑상선 암으로 인해 사망하였다. 55세 이상 환자에서 55세 미만 환자와 비교 시 통계적으로 유의하게 생존 위험도가 증가되어 있었다 (hazard ratio = 35.1, 95% CI = 11.8-104.6, $p < 0.001$, 그림 9). 총 14개의 유전자(*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*)가 55세 이상 환자의 병의 재발 및 생존 위험도 증가와 관련된 유전자로 확인되었다. 각 유전자의 해당 유전자 중 7개(*ENTPD1, HCG11, KBTBD2, PPP2R1B, QSOX2, RINL, ZFAND5*)는 55세 미만의 젊은 환자군에서 발현이 증가되어 있었고, 7개(*ABCF1, F12, FLAD1, GPATCH4, IARS2, NRSN2, RRP36*)는 55세 이상에서 발현이 증가되어 있는 것으로 확인되었다. 각 유전자의 간단한 정보가 보충자료에 제시되어 있다.

그림 9. 나이에 따른 생존 곡선



3. 추출된 유전자와 갑상선 분화도와 상관관계 분석

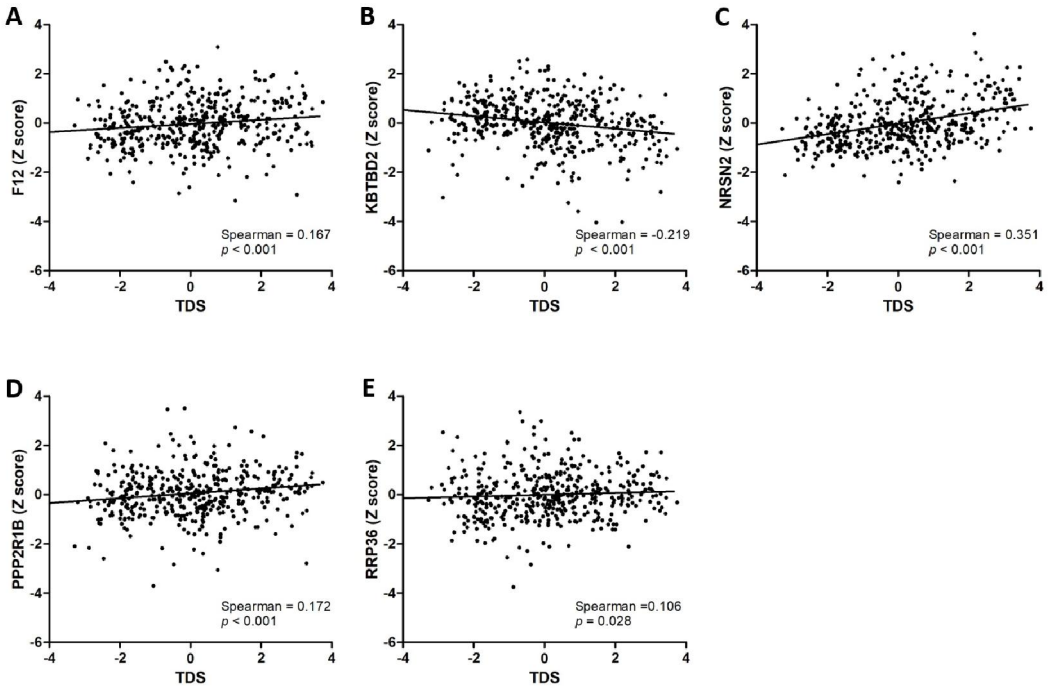
55세 이상 환자에서 나쁜 예후와 관련된 14개의 유전자(*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*)의 발현량(Z 점수)과 갑상선 분화도(TDS, BRS, ERK 점수)의 상관 관계를 분석 하였다. *F12, NRSN2, PPP2R1B, RRP36*은 TDS와 통계적으로 유의하게 양의 상관 관계를 보였으며, *KBTBD2*는 TDS와 통계적으로 유의하게 음의 상관관계를 보였다(표 3, 그림 10). *ABCF1, FLAD1, IARS2, RRP36*은 *BRS*와 통계적으로 유의하게 양의 상관 관계를 보였으며(표 3, 그림 11), *ENTPD1, HCG11, KBTBD2, QSOX2, RINL, ZFAND5*는 *BRS*와 통계적으로 유의하게 음의 상관 관계를 보였다(표 3, 그림 12). *ENTPD1, HCG11, IARS2, KBTBD2, RINL, ZFAND5*는 ERK 점수와 통계적으로 유의하게 양의 상관 관계를 보였으며, *ABCF1, NRSN2*는 ERK 점수와 통계적으로 유의하게 음의 상관 관계를 보였다(표 3, 그림 13).

표 3. 추출된 유전자와 갑상선 분화도와 상관관계 분석 표

mRNA	Highly expressed in	TDS	BRS	ERK score
ABCF1	≥ 55	N/A	+	-
ENTPD1	< 55	N/A	-	+
F12	≥ 55	+	N/A	N/A
FLAD1	≥ 55	N/A	+	N/A
GPATCH4	≥ 55	N/A	N/A	N/A
HCG11	< 55	N/A	-	+
IARS2	≥ 55	N/A	+	+
KBTBD2	< 55	-	-	+
NRSN2	≥ 55	+	N/A	-
PPP2R1B	< 55	+	N/A	N/A
QSOX2	< 55	N/A	-	N/A
RINL	< 55	N/A	-	+
RRP36	≥ 55	+	+	N/A
ZFAND5	< 55	N/A	-	+

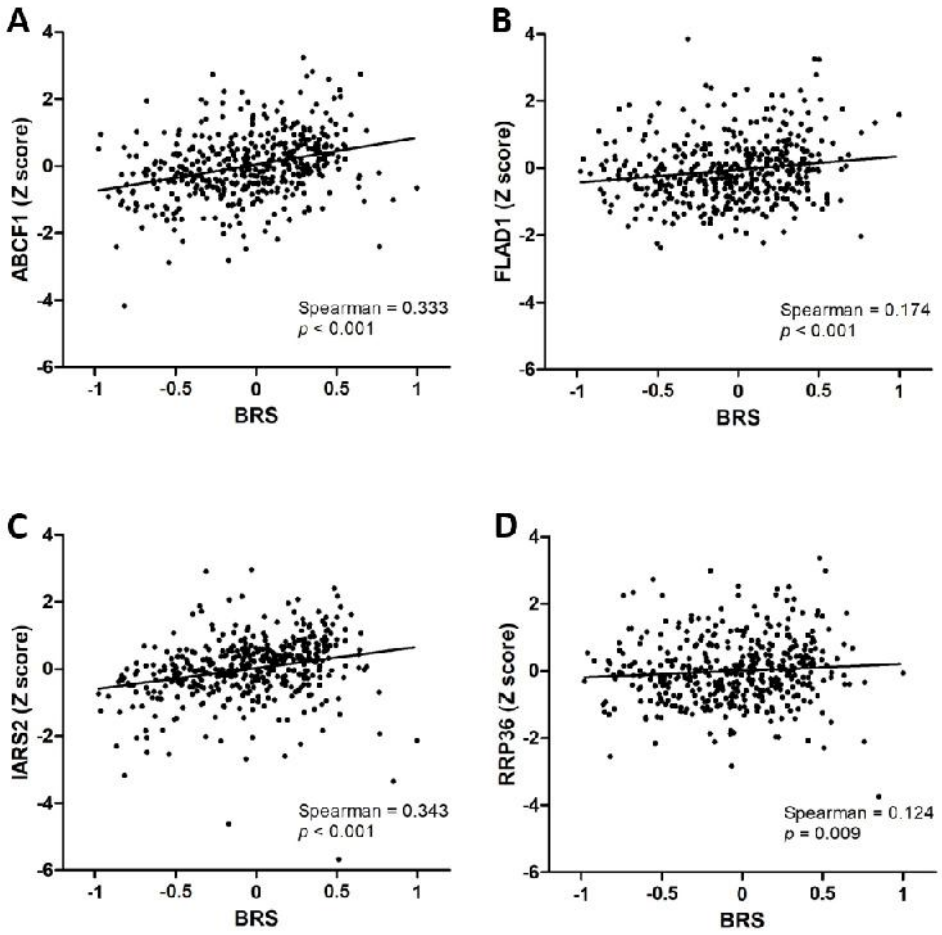
TDS, thyroid differentiation score; BRS, *BRAF*^{V600E}*-RAS* score; ERK, extracellular signal-regulated kinase; N/A., no association

그림 10. Thyroid differentiation score와 유의한 상관 관계를 보이는 유전자 분석



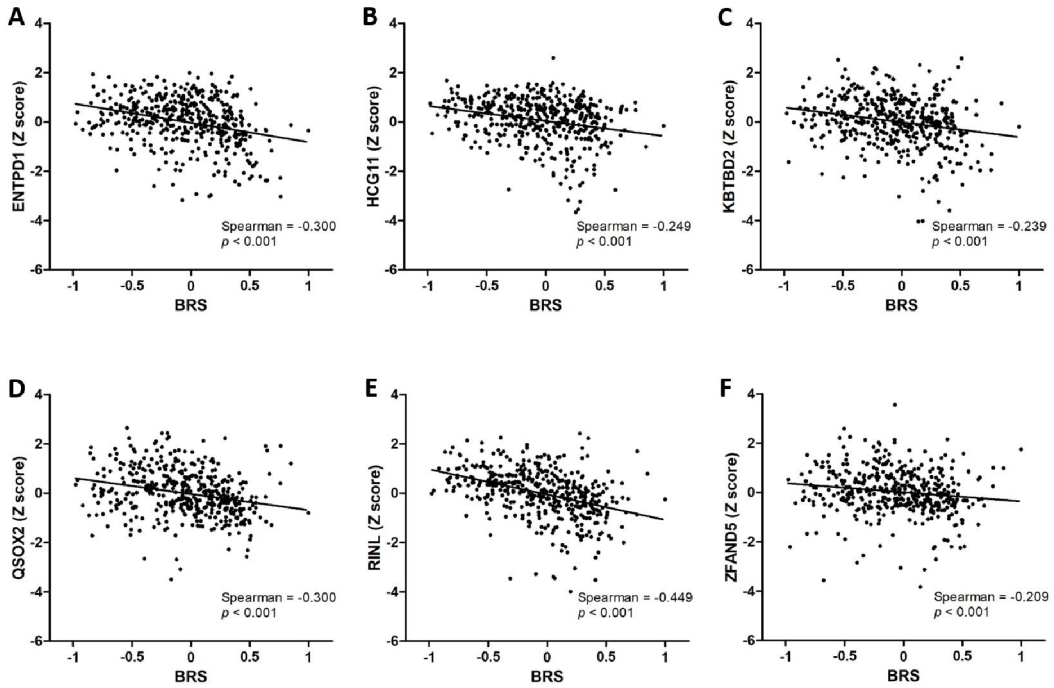
TDS, thyroid differentiation score

그림 11. *BRAF*^{V600E}*-RAS* score 와 유의하게 양의 상관 관계를 보이는 유전자 분석



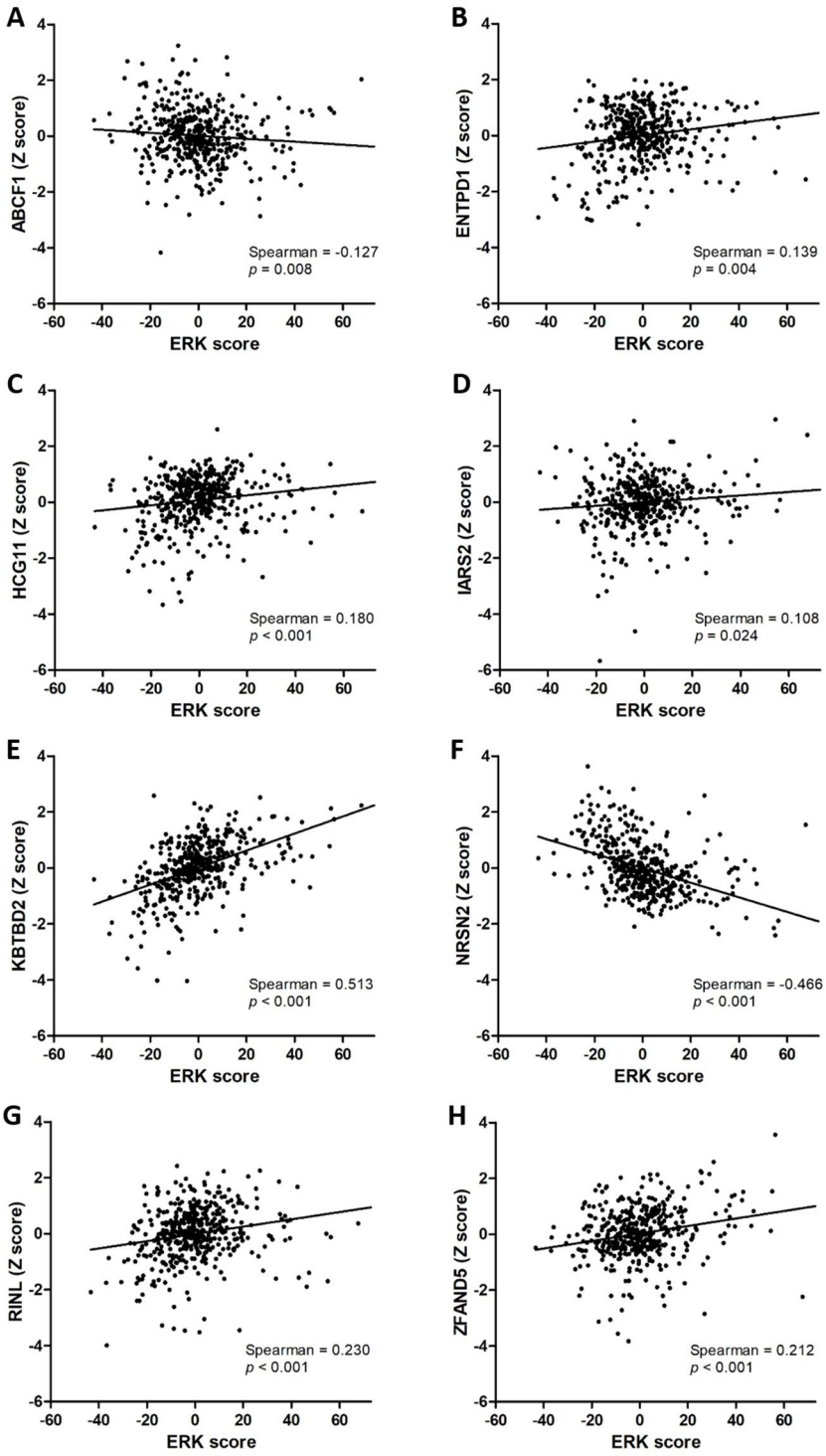
BRS, *BRAF*^{V600E}*-RAS* score

그림 12. *BRAF*^{V600E}*-RAS* score 와 유의하게 음의 상관 관계를 보이는 유전자 분석



BRS, *BRAF*^{V600E}*-RAS* score

그림 13. Extracellular signal-regulated kinase score 와 유의한 상관 관계를 보이는 유전자 분석



ERK, extracellular signal-regulated kinase

고찰

갑상선 암에서 나이가 중요한 예후 인자인 것은 여러 연구에서 잘 알려져 있지만[7-9], 이에 영향을 미치는 유전자에 대한 연구는 부족하다[24,25]. 본 저자들은 Pan-Cancer Atlas 데이터를 이용하여 유두 갑상선 암 환자에서 나이와 관련된 예후에 영향을 미치는 유전자를 확인하였다. 총 14개의 유전자(*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*)가 55세 이상 환자의 나쁜 예후와 관련되는 것을 확인하였으며 각 유전자의 특징을 살펴보고 발현량과 갑상선 분화도와의 상관관계를 분석하였다.

본 연구는 저자들이 아는 한 분화 갑상선 암에서 Pan-Cancer Atlas를 이용해 나이에 따른 예후와 관련된 유전자 분석, 및 해당 유전자와 갑상선 분화도를 계산하여 상관 관계를 확인한 최초 연구이다. 2013년도 암 유전체 아틀라스는 12개 종양에 대하여 종양 게놈의 DNA, 염색질 및 RNA의 이상을 분류하고 후성 유전학과 단백질 결과에 대한 분석 정보를 담고 있다[27]. 2018년도 새롭게 발표된 Pan-Cancer Atlas는 체세포 돌연변이, 영상분석 그리고 종양유전학에서 면역시스템 정보가 포함되어 있으며, mRNA의 분석, 처리, 데이터 분석이 한층 더 발전된 방법을 사용하였다[21,22]. 기존 암 유전체 아틀라스 정보를 활용한 분화 갑상선 암 연구들은 2014년도 발표되었던 연구에서 제시된 갑상선 분화도(TDS, BRS, ERK 점수)를 활용하여 계산한데 반해 본 연구는 Pan-Cancer Atlas를 활용하여 모든 환자들의 갑상선 분화도(TDS, BRS, ERK 점수)를 각각 계산하였고 이를 추출된 유전자들과의 상관 관계를 분석 하였다. 이를 활용한다면 향후에 더 발전된 유전학 정보를 가지고 있는 Pan-Cancer Atlas를 이용하여 분화 갑상선 암의 다양한 유전자 연구가 가능할 것이라고 생각한다.

갑상선 암의 생존에 있어 나이는 가장 중요한 예후 인자 중 하나로, 여

러 갑상선 암 위험도 평가 체계에서 나이가 주요한 요소로 고려된다[1-3]. 하지만 나이에 따른 나쁜 예후에 영향을 미치는 유전자에 대한 연구 결과는 부족하고 잘 알려져 있지 않다[7-9,24,25]. 본 연구는 55세를 기준으로 나이가 많은 환자에서 나쁜 예후에 영향을 미치는 유전자들을 Pan-Cancer Atlas를 활용하여 분석하였고 나이가 갑상선 암에 있어 매우 중요한 예후 인자임을 고려할 때 향후 해당 유전자들을 바탕으로 갑상선 유두 암에 대한 병태생리에 대한 이해와 환자의 진료 및 치료에 도움이 될 수 있을 것이라 생각한다.

추출된 14개의 mRNA 발현량과 BRS 사이의 상관관계를 분석해 보았을 때 4개 유전자(*ABCF1, FLAD1, IARS2, RRP36*)는 BRS와 양의 상관관계를 보이고 이는 *RAS* like 갑상선 유두 암에서 발현이 증가되는 유전자로 판단할 수 있으며, 해당 유전자는 모두 55세 이상에서 발현이 증가되어 있다. 이를 통해 볼 때 해당 유전자(*ABCF1, FLAD1, IARS2, RRP36*)들은 *RAS* like 갑상선 유두 암 환자의 나이에 따른 나쁜 예후에 영향을 미치는 유전자로 생각할 수 있겠다. 또한 6개 유전자(*ENTPD1, HCG11, KBTBD2, QSOX2, RINL, ZFAND5*)는 BRS와 음의 상관관계를 보이고 이는 *BRAF* like 갑상선 유두 암에서 발현이 증가되는 유전자로 판단할 수 있으며, 해당 유전자들은 모두 55세 이상에서 발현이 감소되어 있었다. 이러한 소견을 미루어 볼 때, 6개의 유전자들은 *BRAF* like 갑상선 유두 암 발생의 tumor suppressor와 같은 역할을 하지 않을까 생각된다. *ENTPD1*은 과거 여러 연구들에서 갑상선 암의 좋은 예후와 관련되어 있다는 것이 알려져 있다[28,29](보충자료).

갑상선암의 분화도가 나빠지면 소디움 요오드 동반 수용체(sodium iodide symporter)의 발현이 줄어 방사성요오드 치료에 반응이 나쁘다는 것은 잘 알려진 사실이다[30]. 본 연구에서 총 4개의 유전자(*F12, NRSN2, PPP2R1B, RRP36*)는 TDS와 양의 상관관계를 보였고 이 유전자들은 방사성요오드 치료에 반응이 좋을 갑상선 유두 암을 예측하는데 도움이 될 수 있을 것으로 생각된다. 반면

*KBTBD2*는 TDS와 음의 상관관계를 보이는 유전자로, 방사성요오드 치료에 반응이 좋지 않은 갑상선 유두 암에서 그 발현이 증가되어 있는 유전자로 생각할 수 있겠다.

ERK 점수는 MEK 억제 치료에 발현이 억제 혹은 증가되는 유전자들로 계산된 점수로 높을수록 MAPK 신호 전달 체계가 활성화 되어 있다는 것을 의미하여 ERK 점수는 BRS와 양의 상관관계를 보였다[18,19]. 6개 유전자(*ENTPD1*, *HCG11*, *IARS2*, *KBTBD2*, *RINL*, *ZFAND5*)는 ERK 점수와 양의 상관관계를 보였고, 이들은 MAPK 신호 전달체계가 활성화 되는 종양과 관련되어 있다 판단할 수 있겠다. 그 중, 5개 유전자(*ENTPD1*, *HCG11*, *KBTBD2*, *RINL*, *ZFAND5*)는 BRS점수와는 음의 상관관계를 보였던 유전자들로 *BRAF* like 갑상선 유두 암에서 MAPK 신호전달체계가 활성화되는 것과 관련되겠다. *ABCF1*과 *NRSN2*는 ERK 점수와 음의 상관관계를 보이는 유전자로, 그 발현이 증가되면 MAPK 신호 전달 체계가 억제되는 것으로 생각할 수 있다. *ABCF1*는 BRS와는 양의 상관 관계를 보였던 유전자로 *RAS* like 갑상선 유두 암에서 MAPK 신호 전달체계가 억제되는 것과 관련될 수 있겠다[18,19].

본 연구는 주의 할만한 몇 가지 강점과 제한점이 있다. 본 연구는 분화 갑상선 암의 예후에 영향을 미치는 유전자들을 제시하였다는데 의의가 있다. 또한, 기존에 발표된 암 유전체 아틀라스를 활용한 분화 갑상선암 연구 결과들은 2014년도 발표된 갑상선 분화도(TDS, BRS, ERK 점수)를 활용하여 발표를 하였으나 본 논문에서는 2018년도 발표한 Pan-Cancer Atlas에 제시된 데이터를 이용하여 갑상선 분화도를 직접 계산하여 유전자와 갑상선 분화도 사이 관계를 제시하였다[19,31,32]. 하지만 본 연구는 Pan-Cancer Atlas의 공개 데이터를 이용하여 시행한 연구로 본 논문에서 제시된 유전자로 실제 세포, 실험실 연구를 통해 향후 임상 및 치료에 적용하는 추가적인 연구가 필요하겠다. 또한 본 데이터는 다양한 인종(백인 92.2%, 흑인 4.1%, 아시아인 1.4%, 미국계 인도인 0.2%, 정보 없

는 환자 2.1%)이 포함된 연구로 연구에서 확인된 유전자들이 한국인에게 같은 영향을 미칠지에 대한 추가적인 연구가 필요하겠다.

결론

본 연구는 cBioPortal을 이용하여 2018년도 발표된 Pan-Cancer Atlas 데이터를 통해 갑상선 유두 암 환자에서 나이에 따른 병의 재발 및 생존 위험도 증가와 관련된 14개의 유전자(*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*)를 발견하였다. 또한 이들과 갑상선 분화도와 상관성을 입증하였다. 이를 통해 향후 해당 유전자들을 활용하여 치료 및 예후 평가에 활용 할 수 있는 임상 진료에 활용할 수 있는 후속 연구가 추가적으로 필요하겠다.

참고문헌

1 Amin MB, E. S., Greene F, Byrd DR, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, et al. (Eds.). *AJCC Cancer Staging Manual*. Eight edition edn, (2017).

2 Byar, D. P. *et al.* A prognostic index for thyroid carcinoma. A study of the E.O.R.T.C. thyroid cancer cooperative group. *European Journal of Cancer (1965)* **15**, 1033-1041, doi:https://doi.org/10.1016/0014-2964(79)90291-3 (1979).

3 Edge, S. B. & Compton, C. C. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* **17**, 1471-1474, doi:10.1245/s10434-010-0985-4 (2010).

4 Tuttle, R. M., Haugen, B. & Perrier, N. D. Updated American Joint Committee on Cancer/Tumor-Node-Metastasis Staging System for Differentiated and Anaplastic Thyroid Cancer (Eighth Edition): What Changed and Why? *Thyroid* **27**, 751-756, doi:10.1089/thy.2017.0102 (2017).

5 Kim, M. *et al.* Optimal cut-off age in the TNM Staging system of differentiated thyroid cancer: is 55 years better than 45 years? *Clin Endocrinol (Oxf)* **86**, 438-443, doi:10.1111/cen.13254 (2017).

6 Kim, M. *et al.* Comparison of the Seventh and Eighth Editions of the American Joint Committee on Cancer/Union for International Cancer Control Tumor-Node-Metastasis Staging System for Differentiated Thyroid Cancer. *Thyroid* **27**, 1149-1155, doi:10.1089/thy.2017.0050 (2017).

7 Ganly, I. *et al.* Survival from Differentiated Thyroid Cancer: What Has Age Got to Do with It? *Thyroid* **25**, 1106-1114, doi:10.1089/thy.2015.0104 (2015).

8 Jonklaas, J. *et al.* The impact of age and gender on papillary thyroid cancer s

urvival. *J Clin Endocrinol Metab* **97**, E878-887, doi:10.1210/jc.2011-2864 (2012).

9 Adam, M. A. *et al.* Exploring the Relationship Between Patient Age and Cancer -Specific Survival in Papillary Thyroid Cancer: Rethinking Current Staging Syste ms. *J Clin Oncol* **34**, 4415-4420, doi:10.1200/jco.2016.68.9372 (2016).

10 Nixon, I. J. *et al.* An International Multi-Institutional Validation of Age 55 Years as a Cutoff for Risk Stratification in the AJCC/UICC Staging System for Well-Differentiated Thyroid Cancer. *Thyroid* **26**, 373-380, doi:10.1089/thy.2015.0315 (2016).

11 Chen, A. Y., Jemal, A. & Ward, E. M. Increasing incidence of differentiated thyr oid cancer in the United States, 1988-2005. *Cancer* **115**, 3801-3807, doi:10.100 2/cncr.24416 (2009).

12 Kimura, E. T. *et al.* High prevalence of BRAF mutations in thyroid cancer: gene tic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pat hway in papillary thyroid carcinoma. *Cancer Res* **63**, 1454-1457 (2003).

13 Lemoine, N. R. *et al.* Activated ras oncogenes in human thyroid cancers. *Canc er Res* **48**, 4459-4463 (1988).

14 Pierotti, M. A. *et al.* Rearrangements of TRK proto-oncogene in papillary thyroi d carcinomas. *J Endocrinol Invest* **18**, 130-133, doi:10.1007/bf03349721 (1995).

15 Grieco, M. *et al.* PTC is a novel rearranged form of the ret proto-oncogene a nd is frequently detected in vivo in human thyroid papillary carcinomas. *Cell* **60**, 557-563, doi:10.1016/0092-8674(90)90659-3 (1990).

16 Zaballos, M. A. & Santisteban, P. Key signaling pathways in thyroid cancer. *J E ndocrinol* **235**, R43-r61, doi:10.1530/joe-17-0266 (2017).

17 Agrawal, N. *et al.* Integrated Genomic Characterization of Papillary Thyroid Car cinoma. *Cell* **159**, 676-690, doi:https://doi.org/10.1016/j.cell.2014.09.050 (2014).

18 Pratilas, C. A. *et al.* (V600E)BRAF is associated with disabled feedback inhibitio n of RAF-MEK signaling and elevated transcriptional output of the pathway. *Pr oc Natl Acad Sci U S A* **106**, 4519-4524, doi:10.1073/pnas.0900780106 (2009).

19 Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **159**, 6 76-690, doi:10.1016/j.cell.2014.09.050 (2014).

20 Campbell, P. J. *et al.* Pan-cancer analysis of whole genomes. *Nature* **578**, 82-9 3, doi:10.1038/s41586-020-1969-6 (2020).

21 Ding, L. *et al.* Perspective on Oncogenic Processes at the End of the Beginnin g of Cancer Genomics. *Cell* **173**, 305-320.e310, doi:10.1016/j.cell.2018.03.033 (2 018).

22 Gao, G. F. *et al.* Before and After: Comparison of Legacy and Harmonized TCG A Genomic Data Commons' Data. *Cell Syst* **9**, 24-34.e10, doi:10.1016/j.cels.201 9.06.006 (2019).

23 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for explo ring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10. 1158/2159-8290.CD-12-0095 (2012).

24 Liu, R. & Xing, M. TERT promoter mutations in thyroid cancer. *Endocr Relat C ancer* **23**, R143-155, doi:10.1530/erc-15-0533 (2016).

25 Shen, X. *et al.* Patient Age-Associated Mortality Risk Is Differentiated by BRAF V600E Status in Papillary Thyroid Cancer. *J Clin Oncol* **36**, 438-445, doi:10.120 0/jco.2017.74.5497 (2018).

26 Song, Y. S. & Park, Y. J. Genomic Characterization of Differentiated Thyroid Car cinoma. *Endocrinol Metab (Seoul)* **34**, 1-10, doi:10.3803/EnM.2019.34.1.1 (2019).

27 Cancer Genome Atlas Research, N. *et al.* The Cancer Genome Atlas Pan-Cance r analysis project. *Nat Genet* **45**, 1113-1120, doi:10.1038/ng.2764 (2013).

28 Luo, J.-h., Zhu, Y.-h. & Xiang, C. Favorable function of Ectonucleoside triphosphate diphosphohydrolase 1 high expression in thyroid carcinoma. *Hereditas* **158**, 33, doi:10.1186/s41065-021-00198-6 (2021).

29 Fan, R., Dong, L., Li, P., Wang, X. & Chen, X. Integrated bioinformatics analysis and screening of hub genes in papillary thyroid carcinoma. *PLoS One* **16**, e0251962, doi:10.1371/journal.pone.0251962 (2021).

30 Liu, J., Liu, Y., Lin, Y. & Liang, J. Radioactive Iodine-Refractory Differentiated Thyroid Cancer and Redifferentiation Therapy. *Endocrinology and metabolism (Seoul, Korea)* **34**, 215-225, doi:10.3803/EnM.2019.34.3.215 (2019).

31 Yoo, S. K. *et al.* Integrative analysis of genomic and transcriptomic characteristics associated with progression of aggressive thyroid cancer. *Nat Commun* **10**, 2764, doi:10.1038/s41467-019-10680-5 (2019).

32 Jin, M. *et al.* SHMT2 expression as a diagnostic and prognostic marker for thyroid cancer. *Endocr Connect* **10**, 630-636, doi:10.1530/ec-21-0135 (2021).

영문요약

Background: Age is one of the most important factors predicting the prognosis of papillary thyroid carcinoma (PTC), and it is considered as a major factor in many staging systems. However, there is a lack of genetic data predicting the age-associated mortality of PTC.

Methods: Clinical and genetic data of PTC patients (490 patients) were collected from newly published Pan-Cancer Atlas in 2018 using cBioportal. Patients were divided into two groups; age under 55 (325 patients) and age over 55 (165 patients). First, 812 mRNAs with significantly different expression levels between two groups were collected. Second, analyzed progression free survival and overall survival in each mRNA among patients with age over 55 using R studio, and only statistically significant genes were collected. Third, correlation analysis between mRNAs expression levels and thyroid differentiation scores (thyroid differentiation score (TDS), *BRAF*^{V600E}*-RAS* score (BRS), extracellular signal-regulated kinase (ERK) score) were performed.

Results: The mean age of patients was 47.9 years old and 34% (165 patients) were age over 55. *BRAF* mutation was found in 52% (254 patients) and *RAS* mutation was found in 10% (48 patients). The mortality was significantly increased in patients with age over 55 (hazard ratio = 35.1, $p < 0.001$). With analysis of gene expression levels and age associated mortality, 14 genes (*ABCF1, ENTPD1, F12, FLAD1, GPATCH4, HCG11, IARS2, KBTBD2, NRSN2, PPP2R1B, QSOX2, RINL, RRP36, ZFAND5*) were associated with increased risk for tumor progression or mortality in patients with age over 55. We performed correlation analysis between thyroid differentiation scores in each gene; positive correlation with TDS (*F12*, *NRSN2, PPP2R1B, RRP36*), negative correlation with TDS (*KBTDB2*), positive correlation with BRS (*ABCF1, FLAD1, IARS2, RRP36*), negative correlation with BRS (*ENTPD1, HCG11, KBTBD2, QSOX2, RINL, ZFAND5*), positive correlation with ERK score (*ENTPD1,*

HCG11, IARS2, KBTBD2, RINL, ZFAND5), negative correlation with ERK score (*ABCF1, NRSN2*).

Conclusion: We identified 14 genes predicting the age-associated mortality in PTC patients using Pan-Cancer Atlas data via cBioportal. We also performed correlation analysis between calculated thyroid differentiation scores and each gene. Further study in utilizing genes in clinical practice is needed.

Key words: papillary thyroid carcinoma, Pan-Cancer Atlas, cBioPortal, age, gene analysis

보충자료: 14개 유전자의 간단한 설명[1)]

1. ABCF1

1) Protein: ATP Binding Cassette Subfamily F Member 1
2) Position: 6p21.33
3) Tissue specificity*: enhanced at skeletal muscle cells
4) Single cell type specificity**: low
5) Predicted location: intracellular
6) Molecular function: activator
7) Protein function: Isoform 2 is required for efficient Cap- and IRES-mediated mRNA translation initiation. Isoform 2 is not involved in the ribosome biogenesis.
8) Cancer prognostic summary: liver cancer (unfavorable)
9) Cancer specificity (RNA): low cancer specificity
10) Cancer distribution (RNA): detected in all

2. ENTPD1

1) Protein: Ectonucleoside Triphosphate Diphosphohydrolase 1
2) Position: 10p24.1
3) Tissue specificity: enhanced at endometrium
4) Single cell type specificity: cell type enhanced (macrophage, kupffer cell, monocytes, endothelial cell, B cell)

5) Predicted location: membrane

6) Molecular function: hydrolase

7) Protein function: In the nervous system, could hydrolyze ATP and other nucleotides to regulated purinergic neurotransmission. Could also be implicated in the prevention of platelet aggregation by hydrolyzing platelet-activating ADP to AMP. Hydrolase ATP and ADP equally well.

8) Cancer prognostic summary: renal cancer (unfavorable), thyroid cancer (favorable)

9) Cancer specificity (RNA): low cancer specificity

10) Cancer distribution (RNA): detected in all

3. F12

1) Protein: coagulation factor XII Position: 10p24.1

2) Position: 5p35.3

3) Tissue specificity: enriched at liver

4) Single cell type specificity: cell type enhanced (hepatocyte)

5) Predicted location: secreted

6) Molecular function: hydrolase, protease, serine protease

7) Protein function: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor XIIa and then trypsin cleaves it to beta-factor XIIa. Alpha-factor XIIa activates factor XI to factor XIa.

8) Cancer prognostic summary: renal cancer (unfavorable), liver cancer (favorable)

9) Cancer specificity (RNA): cancer enriched (liver cancer)

10) Cancer distribution (RNA): detected in many

4. FLAD1

1) Protein: Flavin Adenine Dinucleotide Synthetase1

2) Position: 1q21.3

3) Tissue specificity: low tissue specificity

4) Single cell type specificity: low cell type specificity

5) Predicted location: intracellular

6) Molecular function: nucleotidyltransferase, transferase

7) Protein function: Catalyzes the adenylation of flavin mononucleotide (FMN) to form flavin adenine dinucleotide (FAD) coenzyme.

8) Cancer prognostic summary: renal cancer (unfavorable), liver cancer (unfavorable)

9) Cancer specificity (RNA): low cancer specificity

10) Cancer distribution (RNA): detected in all

5. GPATCH4

1) Protein: G-Patch Domain Containing 4

2) Position: 1q22-q23.1

3) Tissue specificity: low tissue specificity

4) Single cell type specificity: low cell type specificity

5) Predicted location: intracellular

6) Cancer prognostic summary: renal cancer (unfavorable), liver cancer (unfavorable)

7) Cancer specificity (RNA): low cancer specificity

8) Cancer distribution (RNA): detected in all

6. HCG11

1) Protein: HLA complex group 11

2) Position: 6p22.2

3) Associated cancer: highly expressed in liver cancer[2)], suppress the growth of glioma[3)], aggravates osteosarcoma carcinogenesis[4)], downregulated in cervical cancer[5)], tumor suppressor in gastric cancer[6)]

7. IARS2

1) Protein: Isoleucyl-TRNA synthetase 2

2) Position: 1q41

3) Tissue specificity: low tissue specificity

4) Single cell type specificity: low cell type specificity

5) Predicted location: intracellular

6) Molecular function: Aminoacyl-tRNA synthetase, Ligase

7) Cancer prognostic summary: pancreatic cancer (unfavorable), cervical cancer (unfavorable)

8) Cancer specificity (RNA): low cancer specificity

9) Cancer distribution (RNA): detected in all

8. KBTBD2

1) Protein: Kelch repeat and BTB domain containing 2
2) Position: 7p14.3
3) Tissue specificity: low tissue specificity
4) Single cell type specificity: low cell type specificity
5) Predicted location: intracellular
6) Cancer prognostic summary: cervical cancer (unfavorable), liver cancer (unfavorable)
7) Cancer specificity (RNA): low cancer specificity
8) Cancer distribution (RNA): detected in all

9. NRSN2

1) Protein: Neurensin 2
2) Position: 20p13
3) Tissue specificity: low tissue specificity
4) Single cell type specificity: cell type enhanced (Intestinal endocrine cells)
5) Predicted location: intracellular, membrane (different isoforms)
6) Protein function: May play a role in maintenance and/or transport of vesicles.
7) Cancer prognostic summary: pancreatic cancer (favorable)
8) Cancer specificity (RNA): low cancer specificity

9) Cancer distribution (RNA): detected in all

10. PPP2R1B

1) Protein: Protein Phosphatase 2 Scaffold Subunit Abeta
2) Position: 11q23.1
3) Tissue specificity: enhanced at liver
4) Single cell type specificity: group enriched (early spermatids, late spermatids, spermatocytes, hepatocyte)
5) Predicted location: intracellular
6) Protein function: The PR65 subunit of protein phosphatase 2A serves as a scaffolding molecule to coordinate the assembly of the catalytic subunit and a variable regulatory B subunit.
7) Cancer prognostic summary: liver cancer (favorable)
8) Cancer specificity (RNA): low cancer specificity
9) Cancer distribution (RNA): detected in many

11. QSOX2

1) Protein: Quiescin Sulfhydryl Oxidase 2
2) Position: 9q34.3
3) Tissue specificity: low tissue specificity
4) Single cell type specificity: cell type enhanced (spermatogonia)
5) Predicted location: intracellular, membrane (different isoforms)

6) Molecular function: oxidoreductase

7) Protein function: Catalyzes the oxidation of sulfhydryl groups in peptide and protein thiols to disulfides with the reduction of oxygen to hydrogen peroxide. May contribute to disulfide bond formation in a variety of secreted proteins. Also seems to play a role in regulating the sensitization of neuroblastoma cells for interferon-gamma-induced apoptosis.

8) Cancer prognostic summary: renal cancer (unfavorable)

9) Cancer specificity (RNA): low cancer specificity

10) Cancer distribution (RNA): detected in all

12. RINL

1) Protein: Ras and Rab Interactor Like

2) Position: 19q13.2

3) Tissue specificity: low tissue specificity

4) Single cell type specificity: cell type enhanced (T-cells)

5) Predicted location: intracellular

6) Molecular function: GTPase activation

7) Protein function: Guanine nucleotide exchange factor (GEF) for RAB5A and RAB22A that activates RAB5A and RAB22A by exchanging bound GDP for free GTP. Plays a role in endocytosis via its role in activating Rab family members (By similarity).

8) Cancer prognostic summary: cervical cancer (favorable), head and neck cancer (favorable)

9) Cancer specificity (RNA): low cancer specificity

10) Cancer distribution (RNA): detected in all

13. RRP36

1) Protein: Ribosomal RNA Processing 36

2) Position: 6p21.1

3) Tissue specificity: low tissue specificity

4) Single cell type specificity: low cell type specificity

5) Predicted location: intracellular

6) Protein function: Involved in the early processing steps of the pre-rRNA in the maturation pathway leading to the 18S rRNA.

7) Cancer prognostic summary: renal cancer (unfavorable), liver cancer (unfavorable)

8) Cancer specificity (RNA): low cancer specificity

9) Cancer distribution (RNA): detected in all

14. ZFAND5

1) Protein: Zinc Finger AN1-Type Containing 5

2) Position: 9q21.13

3) Tissue specificity: tissue enhanced (skeletal muscle)

4) Single cell type specificity: low cell type specificity

5) Predicted location: intracellular

6) Protein function: Involved in protein degradation via the ubiquitin-proteasome system. May act by anchoring ubiquitinated proteins to the proteasome. Plays a role in ubiquitin-mediated protein degradation during muscle atrophy. Plays a role in the regulation of NF-kappa-B activation and apoptosis. Inhibits NF-kappa-B activation triggered by overexpression of RIPK1 and TRAF6 but not of RELA. Inhibits also tumor necrosis factor (TNF), IL-1 and TLR4-induced NF-kappa-B activation in a dose-dependent manner. Overexpression sensitizes cells to TNF-induced apoptosis. Is a potent inhibitory factor for osteoclast differentiation.

7) Cancer prognostic summary: endometrial cancer (favorable)

8) Cancer specificity (RNA): low cancer specificity

9) Cancer distribution (RNA): detected in all

* The RNA specificity category is based on mRNA expression levels in the analyzed samples based on a combination of data from HPA, GTEX and FANTOM5. The categories include: tissue enriched, group enriched, tissue enhanced, low tissue specificity and not detected.

** The RNA specificity category is based on mRNA expression levels in the analyzed cell types based on scRNA-seq data from normal tissues. The categories include: cell type enriched, group enriched, cell type enhanced, low cell type specificity and not detected.

Reference

1. Available from: www.proteinatlas.org.

2. Xu Y, Zheng Y, Liu H, Li T. Modulation of IGF2BP1 by long non-coding RNA HCG11 suppresses apoptosis of hepatocellular carcinoma cells via MAPK signal ing transduction. International journal of oncology 2017;51:791-800.

3. Zhang L, Cao Y, Kou X, Che L, Zhou X, Chen G, et al. Long non-coding RNA HCG11 suppresses the growth of glioma by cooperating with the miR-4425/MTA3 axis. J Gene Med 2019;21:e3074.
4. Wang L, Zhou J, Zhang Y, Hu T, Sun Y. Long Non-Coding RNA HCG11 Aggravates Osteosarcoma Carcinogenesis via Regulating the microRNA-579/MMP13 Axis. Int J Gen Med 2020;13:1685-95.
5. Zhang Y, Zhang J, Mao L, Li X. Long noncoding RNA HCG11 inhibited growth and invasion in cervical cancer by sponging miR-942-5p and targeting GFI1. Cancer Med 2020;9:7062-71.
6. Zhang Q, Yang K, Li J, Chen F, Li Y, Lin Q. Long Noncoding RNA HCG11 Acts as a Tumor Suppressor in Gastric Cancer by Regulating miR-942-5p/BRMS1 Axis. Journal of Oncology 2021;2021:9961189.