의학박사 학위논문

Integrative Analysis on Malignant Pheochromocytoma or Paraganglioma in TCGA and EMBL-EBI

TCGA와 EMBL-EBI 분석에 의한 악성 갈색세포종 혹은 악성 부신경절종 연구

2018년 2월

서울대학교 대학원

의학과 외과학

서 용 준

Integrative Analysis on Malignant Pheochromocytoma or Paraganglioma in TCGA and EMBL-EBI

지도교수 김 수 진

이 논문을 의학박사 학위논문으로 제출함
2017년 10월

서울대학교 대학원
의학과 외과학
서 용 준

서용준의 의학박사 학위논문을 인준함
2017년 12월

위 원 장 _____ (인)
부위원장 _____ (인)
위 원 _____ (인)
위 원 _____ (인)
위 원 _____ (인)

Abstract

Integrative Analysis on Malignant Pheochromocytoma or Paraganglioma in TCGA and EMBL-EBI

Yong Joon Suh

Department of Medicine (Surgery)

The Graduate School

Seoul National University

Introduction: Methods of predicting malignant behavior of pheochromocytomas (PCCs) or paragangliomas (PGLs) are needed. However, there are few reliable histopathologic criteria to predict malignant behavior in PCC/PGLs. Recent genomic analysis of The Cancer Genome Atlas (TCGA) and European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) has provided genetic information enabling more accurate differentiation of disease entities. TCGA and EMBL-EBI could be utilized for the genetic analysis of malignant PCC/PGLs. Therefore, the objective of this study was to determine genomic expression differences and mutational differences of malignant PCC/PGL using data from TCGA and EMBL-EBI.

Methods: TCGA had data of multigenomic analysis for 184 PCC/PGL samples while EBML-EBI had data for 202 PCC/PGL samples. Clinical information, mutation status, and mRNA expression dataset were downloaded from TCGA and EMBL-EBI. Following inclusion/exclusion criteria, 59 of 184 PCC/PGL samples in TCGA and 171 of 202 PCC/PGL samples in EMBL-EBI were selected. Of 59 samples in TCGA, 12 were malignant and 47 were benign. Of 171 samples in EMBL-EBI, 19 were malignant and 152 were benign. Data

of mRNA expression and mutations were compared between these two groups.

Results: Thirty up/down-regulated pathways in malignant PCC/PGLs were related to cancer signaling, metabolic alteration, prominent mitosis, and junctional dissociation. Twenty-one up-regulated genes and 11 down-regulated genes were significantly enriched in functional annotation pathways. Malignancy was more often in germline mutations of *SDHB* or *VHL*. Four candidate mutational genes (*ATRX, SETD2, MUC16*, and *PAX1*) were proposed as genes susceptible to malignancy. Overall survival rates were significantly correlated with the presence or absence of these somatic mutations.

Conclusions: TCGA and EMBL-EBI showed differences in mRNA expression and mutations between malignant and benign PCC/PGLs. Improved recognition of prognostic factors can help us achieve proper diagnosis and provide appropriate treatment for PCC/PGLs.

Keywords: Malignancy, Pheochromocytoma, Paraganglioma, TCGA, EMBL-EBI

Student Number: 2015-30562

목 차

List of Tables

List of Figures

Introduction

Pheochromocytomas (PCCs) are catecholamine-producing neuroendocrine neoplasms originating from adrenal medulla [1]. Paragangliomas (PGLs) are catecholamine-secreting neuroendocrine neoplasms originating from extra-adrenal chromaffin tissues of sympathetic ganglia. Only 15–20% of PCC/PGLs arise from extra-adrenal chromaffin tissues whereas 80–85% of PCC/PGLs develop from adrenal medulla. Annual incidence of PCC/PGLs is between 2 and 8 per million. The prevalence of PCC/PGLs in the population is 1:6,500 to 1:2,500, respectively [2, 3]. These neoplasms present with headache, diaphoresis, tachycardia, hypertension, and anxiety [4]. The symptoms are due to the release of catecholamines.

Diagnostic tests for PCC/PGLs include biochemical evaluation, imaging, and histopathology, in addition to genetic testing [5-7]. Germline mutations have been found in 20-41% cases of PCC/PGLs [3]. Commercial panel tests include 10 gene mutations: *MAX*, *MEN1*, *NF1*, *RET*, *SDHAF2*, *SDHB*, *SDHC*, *SDHD*, *TMEM127*, and *VHL*. For management of patients with PCC/PGLs, physicians can consider treatments such as medical management, surgery, targeted radiation therapy of ^{131}I-MIBG, chemotherapy, embolization, cryoablation, radiofrequency ablation, and targeted molecular therapy [8]. Since there is no curative management for malignant PCC/PGLs, managements usually focus on vigilant monitoring of metastasis in malignant PCC/PGLs.

Considering the rule of 10s, 10% of PCC/PGLs are considered malignant [9, 10]. However, the malignancy rate certainly exceeds 10% in patients with extra-adrenal disease [11]. According to one study, 5% of adrenal-based PCCs and 33% of extra-adrenal PGLs are malignant [1]. Malignant PCC/PGLs show 5-year survival rates of around 50% [12, 13]. Their long-term survival has not been improved over the last two decades [14]. The exact prevalence of malignant PCC/PGLs is difficult to verify because these neoplasms can metastasize after resection of a mass regarded as benign.

Histologic analysis alone cannot predict the malignant or benign behavior of PCC/PGLs [15]. According to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system, the distinctive feature of malignant PCC/PGLs is only distant metastasis. Therefore, their diagnostic limitations are impeding therapeutic planning. Proper diagnosis is directly linked to proper management and follow-up according to disease prognosis. Institutions have strived to define molecular markers for malignant PCC/PGLs [16, 17]. Efforts are ongoing to identify solid predictors of malignancy [18, 19].

In patients with PCC/PGLs, genetic counseling is commonly recommended owing to the risk of the disease itself and the importance of hereditary factors in the development of the disease [20-22]. Molecular diagnosis is intrinsically more essential than microscopic diagnosis. In the near future, genotype will be one of objective aspects used to stratify patients with PCC/PGLs and determine a therapeutic approach.

Currently, molecular characterization data from The Cancer Genome Atlas (TCGA) and European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) are available to the public [23-25]. These public datasets describe tumors including their genomic sequence, mutation, DNA methylation, copy-number variation, mRNA expression, microRNA expression, reverse phase protein array, and clinical information. TCGA and EMBL-EBI could be utilized for genetic analysis of malignant PCC/PGLs.

The aim of the present study was to investigate genomic and mutational differences of malignant PCC/PGLs based on data from TCGA and EMBL-EBI.

Materials and Methods

Study design

Accessible data from TCGA and EMBL-EBI were obtained to compare malignant and benign PCC/PGLs. TCGA had multigenomic analysis data of 184 PCC/PGL samples. Initial pathologic diagnosis was made from 1988 to 2013. Clinical information, mutation status, and gene expression dataset (20,531 genes) of RNA-sequencing mRNA Fragments Per Kilobase Million (FPKM) were downloaded from TCGA. The dataset was integrated into a table. Only samples with TCGA type code 01 (primary solid tumor) were included (Figure 1). Two samples with type code 05 from additional primary sites and 1 sample with type code 06 from a metastatic site were excluded. Benign samples were included only if the follow-up was more than two years. However, malignant samples were included even if they were metastasized within two years. Aggressive samples with local invasion or metastatic lymph nodes were excluded due to uncertain behavior. 59 samples met the inclusion criteria. Among the 59 PCC/PGL samples, 12 were malignant and 47 were benign. Clinical characteristics, mutations, and mRNA expression were reorganized according to behavior and compared to each other. EMBL-EBI had multi-omics analysis data of 202 PCC/PGL samples from the French 'Cortico et Médullosurrénale: les Tumeurs Endocrines' (COMETE) Network. Cases were recruited from 1993 to 2008. Clinical information, mutation status, and expression dataset (39,534 probes) of RNA-sequencing mRNA transcripts were downloaded from EMBL-EBI. Only primary tumor samples were included. The selection criteria excluded the following: 1) 13 samples without available data, 2) 7 duplicated samples from same sources, 3) 1 sample from psoas muscle, 4) 4 samples from lymph nodes, and 5) 4 double primary samples with both PCC and PGL (Figure 2). Even in eligible samples, 2 samples were further excluded because they were not labeled as malignant or benign. A total of 171 samples had accessible genomic data and corresponding clinical information. Among the 171 samples, 19

were malignant whereas 152 were benign. Clinical characteristics, mutations, and mRNA expression were reorganized according to behavior and compared to each other. For interpretation of gene expression, differentially expressed genes (DEGs) were extracted from TCGA and EMBL-EBI. Commonly enriched genes were searched by means of functional annotation pathway analysis. TCGA opened whole-exome sequencing somatic mutations. EMBL-EBI disclosed targeted Sanger sequencing somatic mutations although whole-exome sequencing was performed only for 31 pairs. Therefore, pivotal somatic mutations for malignancy were queried from TCGA and those mutations were cross-checked in EMBL-EBI. The information on germline mutations was more restricted in TCGA and EMBL-EBI. All accessible data in both databases were utilized for analysis of germline mutations.

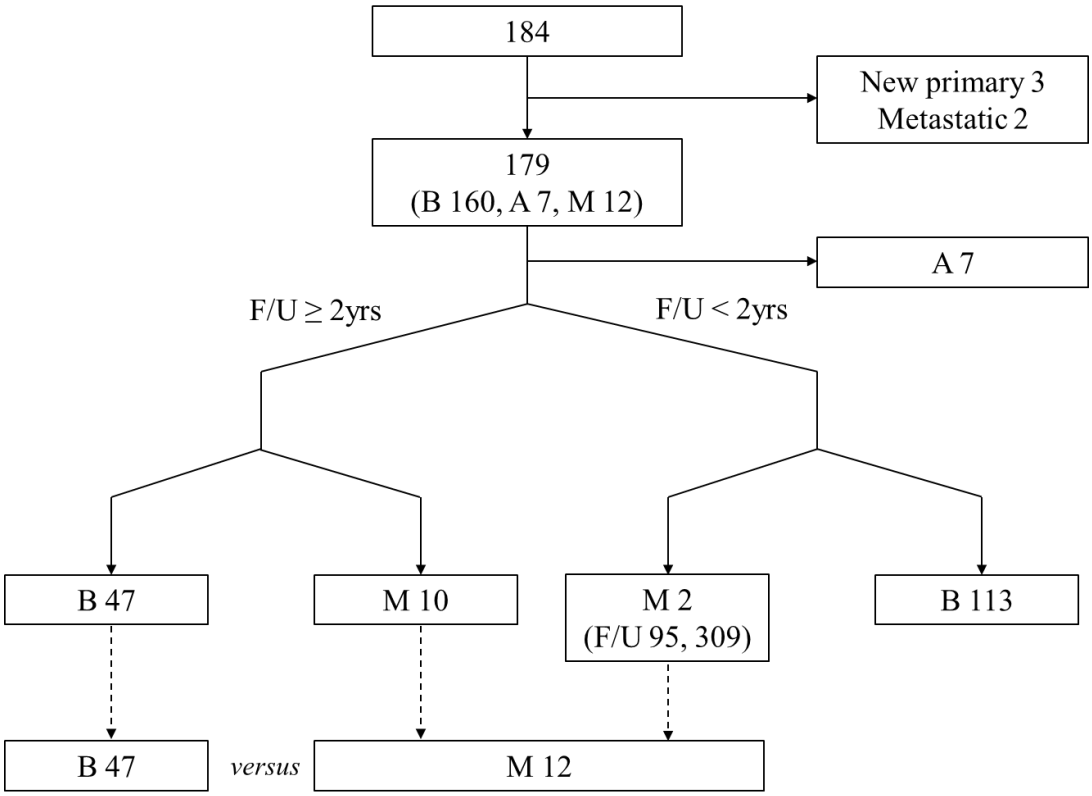


Figure 1. Flow diagram showing the process of enrollment in TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma, *B* benign, *A* aggressive, *M* malignant, *F/U* follow-up

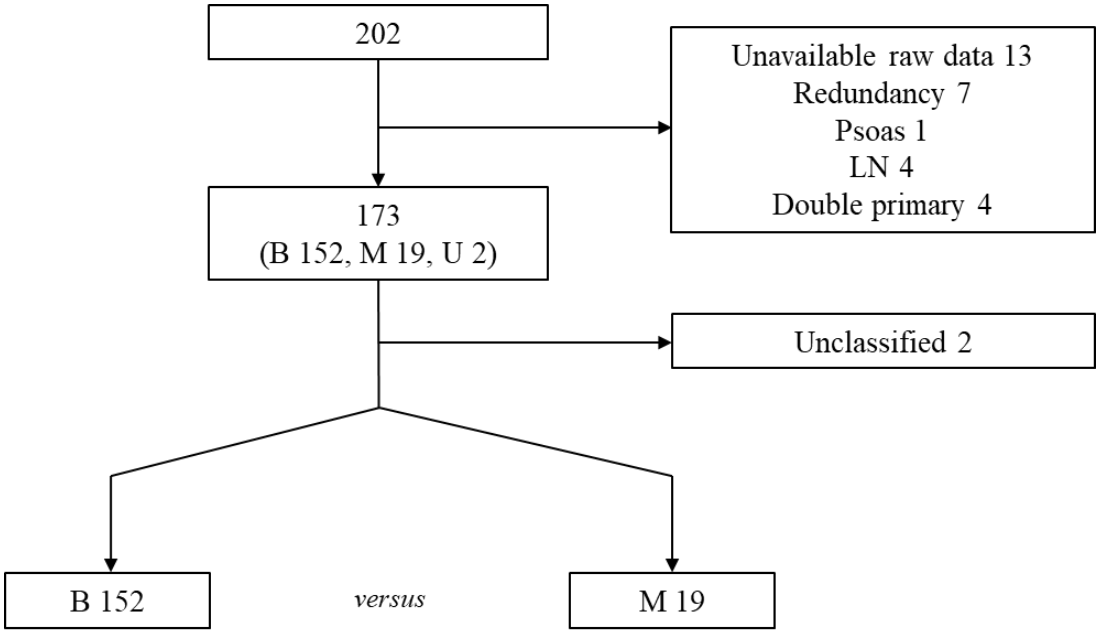


Figure 2. Flow diagram showing the process of enrollment in EMBL-EBI. *PCC* pheochromocytoma, *PGL* paraganglioma, *B* benign, *M* malignant, *U* unclassified, *LN* lymph node

Definition

Metastasis was defined as the appearance of chromaffin tissues at non-chromaffin sites distant from the primary tumor. Malignant or benign behavior was used to describe metastasis or no metastasis, respectively. In TCGA, 12 PCC/PGL samples considered as malignant were acquired from patients with metastasis. In EMBL-EBI, 19 PCC/PGL samples were labeled as malignant. Zero value was filtered for the analysis of mRNA. To identify DEGs, false discovery rate (FDR) <0.05 and $\log_2$ fold change (FC) ≥ 1.5 were set as the thresholds. Functional enrichment of genes was analyzed based on Gene Ontology. Kyoto Encyclopedia of Genes and Genomes (KEGG) was used as reference pathway database.

Statistics

Malignant and benign PCC/PGL groups were compared using Chi-square test or Fisher's exact test for categorical data and Student's t test or Mann–Whitney *U* test for continuous data without normal distribution. Overall survival rates were estimated using Kaplan–Meier analysis, including log-rank test. Overall survival intervals were expressed as days from the date of surgery to the date of the last follow-up. In two-tailed tests, a p value of less than 0.05 was considered statistical significant. Statistical analysis was performed in R 3.4.1 (R Foundation for Statistical Computing, Vienna, Austria) and GSEA 2.1.0 (Broad Institute, Cambridge, MA, USA).

Results

Clinical characteristics

A total of 59 patients with PCC/PGLs from TCGA were included in the current study. Among these patients, 12 with metastasis were assigned into the malignant PCC/PGL group. In the malignant PCC/PGL group, 9 (75.0%) patients showed progressive disease in response to treatment while only one (8.3%) patient experienced complete remission (Table 1). No demographic difference was found between malignant and benign PCC/PGLs (Table 2). The proportion of paraganglioma was significantly ($p=0.010$) higher in malignant PCC/PGLs. Dopamine secretion was more frequent in malignant PCC/PGLs whereas metanephrine secretion was only found in benign PCC/PGLs (Table 3). The size was also significantly greater in malignant PCC/PGLs ($p<0.001$). Their mean follow-up periods were 2423.5 ± 2538.3 days and 1617.6 ± 867.7 days in malignant and benign PCC/PGLs, respectively ($p=0.728$).

Table 1 Treatment responses in 12 PCC/PGL samples with malignant behavior in TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma

Treatment response	Number of samples (n=12)
Complete remission	1
Partial remission	1
Stable disease	1
Progressive disease	9

Table 2 Comparison of identified characteristics between malignant and benign PCC/PGLs from TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma

Characteristic	Malignant (n=12)	Benign (n=47)	p value
Age (years), mean ± SD	43.1 ± 13.9	43.4 ± 13.3	0.903
Gender, n (%)	12	47	
Male	7 (58.3)	24 (51.1)	0.752
Female	5 (41.7)	23 (48.9)	
Race, n (%)	12	46	0.455
White	11 (91.7)	36 (78.2)	
African American	1 (8.3)	5 (10.9)	
Asia	0 (0)	5 (10.9)	
Laterality, n (%)	6	41	0.749
Right	2 (33.3)	19 (46.4)	
Left	4 (66.7)	21 (51.2)	
Bilaterality	0 (0)	1 (2.4)	
Diagnosis, n (%)	12	47	0.010
PCC	6 (50.0)	41 (87.2)	
PGL	6 (50.0)	6 (12.8)	
Size (mm), mean ± SD	81.5 ± 27.6	44.1 ± 14.3	<0.001
Follow-up (days), mean ± SD	2423.5 ± 2538.3	1617.6 ± 867.7	0.728

Table 3 Catecholamine secretion features between malignant and benign PCC/PGLs from TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma

Characteristic	Malignant (n=9)	Benign (n=40)	p value
Biochemical testing			
Normetanephrine	6	34	0.336
Norepinephrine	5	24	1.000
Epinephrine	1	15	0.238
Metanephrine	0	21	0.006
Methoxytyramine	1	0	0.184
Dopamine	4	5	0.046

A total of 171 patients with PCC/PGLs from EMBL-EBI were included for cross-validation. No age or gender difference was found between malignant and benign PCC/PGLs (Table 4). The proportion of paraganglioma was significantly higher in malignant PCC/PGLs ($p<0.001$). The size was also significantly greater in malignant PCC/PGLs ($p<0.004$). A total of 189 samples had information of tumor size, including 48 samples from TCGA and 141 samples from EMBL-EBI. From the 189 samples of TCGA and EMBL-EBI, the optimal cut-off value was calculated to identify malignancy. According to ROC curves, the size of 54.5 mm showed the highest accuracy (AUC=0.778, sensitivity=66.7%, specificity=61.6%, $p<0.001$) (Figure 3).

Table 4 Comparison of identified characteristics between malignant and benign PCC/PGLs from EMBL-EBI. *PCC* pheochromocytoma, *PGL* paraganglioma

Characteristic		Malignant (n=19)	Benign (n=152)	p value
Age (years), mean ± SD		36.4 ± 14.1	43.2 ± 16.4	0.079
Gender, n (%)		19	152	
	Male	9 (47.4)	59 (38.8)	0.780
	Female	10 (52.6)	93 (61.2)	
Diagnosis, n (%)		19	152	<0.001
	PCC	9 (47.4)	136 (89.5)	
	PGL	10 (52.6)	16 (10.5)	
Tumor size (mm), mean ± SD		67.8 ± 28.0	47.3 ± 21.8	0.004

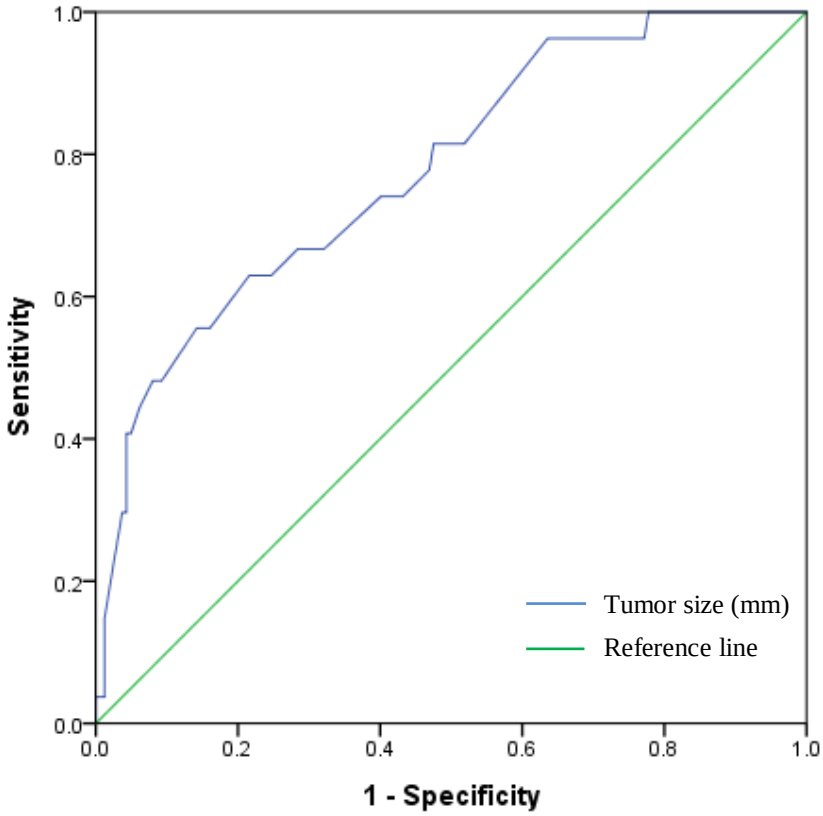
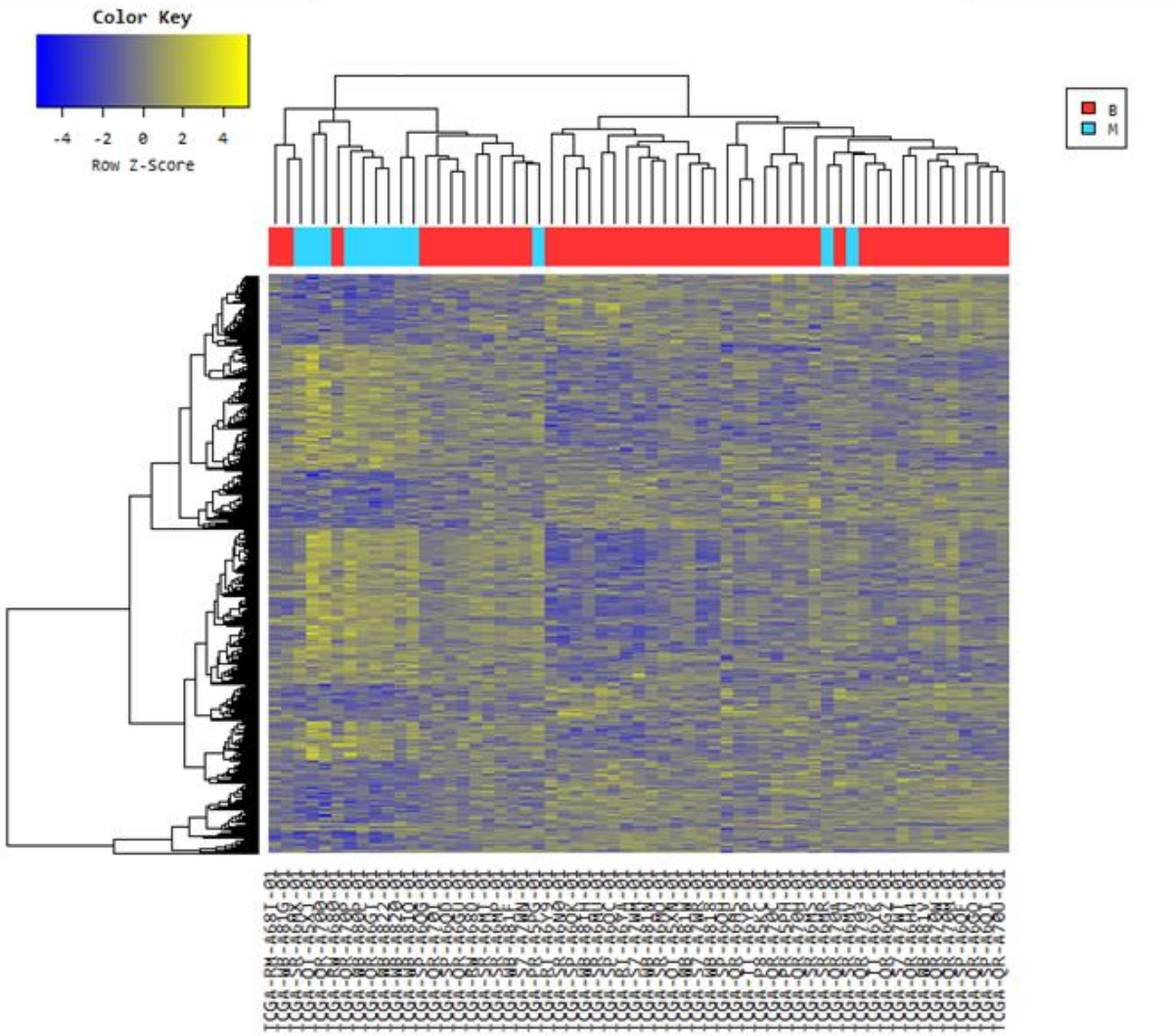


Figure 3. According to ROC curves, the size of 54.5 mm showed the highest accuracy in TCGA and EMBL-EBI (AUC=0.778, sensitivity=66.7%, specificity=61.6%, p<0.001). *ROC* receiver operating characteristic, *AUC* area under the curve

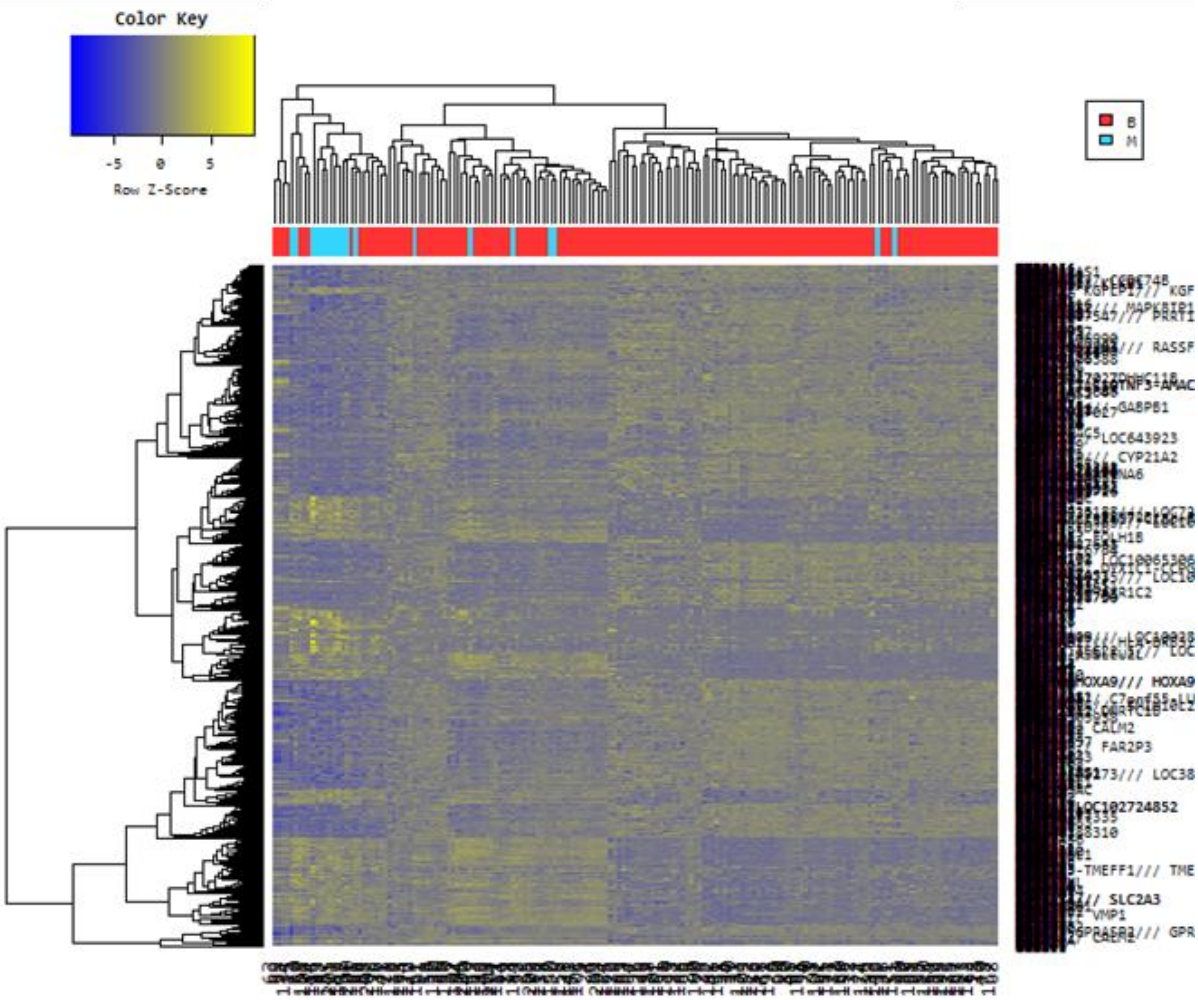
mRNA analysis

In TCGA, 6,056 out of 20,531 genes were excluded because they had zero value in at least one sample. By comparative analysis of 14,474 genes, 367 up-regulated genes were identified while 282 genes were down-regulated. A total of 39,534 probes were used in EBML-EBI. Results showed that 558 probes were up-regulated while 1,132 probes were down-regulated. In both data, quality control was performed for each gene or probe. Distribution of expression level was examined in various ways. Hierarchical clustering analysis was performed by Euclidean distance and complete linkage (Figure 4). Gene enrichment was identified by functional annotation analysis (Figure 5). Commonly enriched pathways in malignant PCC/PGLs were linked to cancer signaling, metabolic alteration, prominent mitosis, and junctional dissociation (Table 5). Thirty up/down-regulated pathways harbored up-regulated genes such as *BIRC5*, *BRIP1*, *BUB1*, *CCNA2*, *CCNB1*, *CDK1*, *DIAPH3*, *ELOVL2*, *ELOVL7*, *ESPL1*, *FANCI*, *IQGAP3*, *ITGAV*, *MKI67*, *PMAIP1*, *RMI2*, *RRM2*, *RYR2*, *TK1*, *TOP2A*, and *TYMS*. They also encompassed down-regulated genes such as *ACOT7*, *CDKN1C*, *GPX3*, *HDAC11*, *HIST1H2AC*, *HIST1H4H, KCNJ5*, *NTNG1*, *PRKACB*, and *RET*. These 32 common genes were investigated for gene annotation (Table 6).



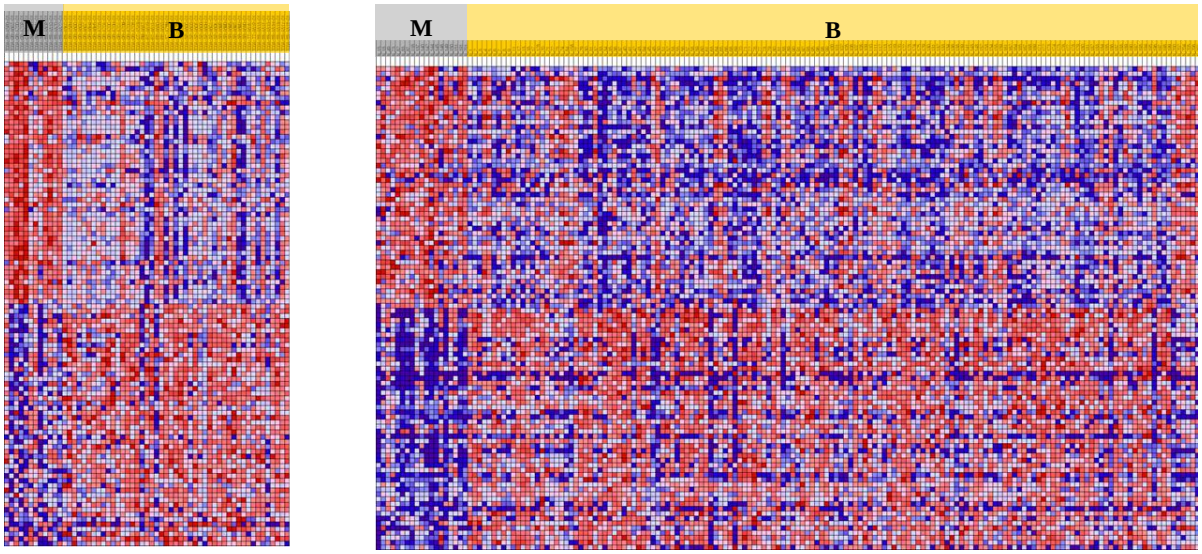
(a)

(*Continued*)



(b)

Figure 4. (*Continued*) Hierarchical clustering analysis performed by Euclidean distance and complete linkage of (a) TCGA and (b) EMBL-EBI. *M* malignant, *B* benign



(a) (b)

Figure 5 Gene Set Enrichment Analysis showing heat map of top 50 features for each phenotype according to the presence of metastasis in PCC/PGLs from (a) TCGA and (b) EMBL-EBI. *PCC* pheochromocytoma, *PGL* paraganglioma, *M* malignant, *B* benign

Table 5 List of 32 common genes up/down-regulated commonly in TCGA and EMBL-EBI.

Genes	Description
BUB1	BUB1 mitotic checkpoint serine/threonine kinase
PMAIP1	phorbol-12-myristate-13-acetate-induced protein 1
RYR2	ryanodine receptor 2
TOP2A	topoisomerase (DNA) II alpha
IQGAP3	IQ motif containing GTPase activating protein 3
DIAPH3	diaphanous related formin 3
BIRC5	baculoviral IAP repeat containing 5
ESPL1	extra spindle pole bodies like 1, separase
CDK1	cyclin dependent kinase 1
RRM2	ribonucleotide reductase regulatory subunit M2
MKI67	marker of proliferation Ki-67
BRIP1	BRCA1 interacting protein C-terminal helicase 1
TK1	thymidine kinase 1
TYMS	thymidylate synthetase
ELOVL7	ELOVL fatty acid elongase 7
CCNA2	cyclin A2
ELOVL2	ELOVL fatty acid elongase 2
RMI2	RecQ mediated genome instability 2
FANCI	Fanconi anemia complementation group I
CCNB1	cyclin B1
ITGAV	integrin subunit alpha V
TSC2	tuberous sclerosis 2
ACOT7	acyl-CoA thioesterase 7
PRKACB	protein kinase cAMP-activated catalytic subunit beta
CDKN1C	cyclin dependent kinase inhibitor 1C
GPX3	glutathione peroxidase 3
HDAC11	histone deacetylase 11
HIST1H4H	histone cluster 1 H4 family member h
HIST1H2AC	histone cluster 1 H2A family member c
KCNJ5	potassium voltage-gated channel subfamily J member 5
NTNG1	netrin G1
RET	ret proto-oncogene

Table 6 List of 30 up/down-regulated pathways between malignant and benign PCC/PGLs from TCGA and EMBL-EBI. *PCC* pheochromocytoma, *PGL* paraganglioma, *GO* gene ontology

GO term	GO function	Gene expression		p value	
		UP	DOWN	TCGA	EMBL-EBI
Metabolism	fatty acid elongation	*ELOVL2, ELOVL7*	*ACOT7*	0.009630	0.000198
	biosynthesis of unsaturated fatty acid	*ELOVL2*	*ACOT7*	0.123855	0.002226
	glutathione metabolism	*RRM2*	*GPX3*	0.003844	0.000325
	pyrimidine metabolism	*RRM2, TK1, TYMS*		<0.000001	0.022632
	purine metabolism	*RRM2*		0.000071	0.000003
Cellular Process	cell cycle	*BUB1, CDK1, CCNA2, ESPL1, MKI67, TOP2A*	*CDKN1C*	<0.000001	0.000009
	p53 signaling pathway	*CCNB1, CDK1, PMAIP1, RRM2*		<0.000001	<0.000001
	apoptosis	*BIRC5, PMAIP1*		0.000152	0.003513
	cellular senescence	*CCNA2, CCNB1, CDK1*	*TSC2*	<0.000001	<0.000001
	regulation of actin cytoskeleton	*DIAPH3, IQGAP3, ITGAV*		0.031759	<0.000001
	tight junction		*PRKACB*	0.002898	0.000003
	gap junction	*CDK1*	*PRKACB*	0.013933	0.000390
Environment Information	cell adhesion molecule	*ITGAV*	*NTNG1*	0.001482	<0.000001
	PI3-AKT signaling pathway	*ITGAV*	*TSC2*	0.443009	<0.000001
	mTOR signaling pathway		*TSC2*	0.205507	0.000006
	Ca^{2+} signaling pathway	*RYR2*		0.000633	<0.000001
	Apelin signaling pathway	*RYR2*	*PRKACB*	0.177755	<0.000001
	cAMP signaling pathway	*RYR2*	*PRKACB*	0.025700	<0.000001
	MAPK signaling pathway		*PRKACB*	0.014752	<0.000001
	RAS signaling pathway		*PRKACB*	0.133572	<0.000001
	WNT signaling pathway		*PRKACB*	0.046904	0.000026
	FOXO signaling pathway	*CCNB1*		0.006865	0.000093
	AMPK signaling pathway	*CCNA2*		0.146879	0.000053
Genetic Information	Fanconi anemia pathway	*BRIP1, FANCI, RMI2*		<0.000001	0.020855
	homologous recombination	*BRIP1*		0.000005	0.069857
Organismal Systems	chemokine signaling pathway		*PRKACB*	0.020752	<0.000001
	insulin secretion	*RYR2*	*PRKACB*	0.012737	0.001032
	circadian entrainment	*RYR2*	*KCNJ5, PRKACB*	0.000246	<0.000001
Human Disease	pathways in cancer	*BIRC5, ITGAV*	*PRKACB, RET*	<0.000001	<0.000001
	alcoholism		*HDAC11, HIST1H2AC, HIST1H4H*	0.000010	<0.000001

Mutation analysis

Based on data from TCGA, 2,632 somatic mutations were identified without considering frequency of mutations. Of these mutations, 1,546 (58.7 %) were missense mutations and 653 (24.8 %) were silent mutations (Figure 6). Silent mutations were discarded for mutational analysis. Four mutations were found to be significantly more frequent in malignant PCC/PGLs. These mutations were *ATRX, SETD2, MUC16,* and *PAX1* (Table 7). These genetic alterations were correlated with personal neoplasm status (Figure 7). These mutations were expected to affect each domain. Furthermore, these mutations were significantly correlated with overall survival rates (Figure 8). Kaplan-Meier curve showed the most significant value when patients with at least one of these mutations were compared to patients with none of these mutations ($p<0.001$) (Figure 9). Intergenic interactions of these genes were visualized in platforms of gene expression, copy number variation, methylation, and protein expression by techniques described previously (Figure 10) [26]. These genes other than *PAX1* worked positively in the same direction. Especially, *ATRX* and *SETD2* had significant tendency towards co-occurrence ($p=0.046$).

Figure 6 Variant classification of somatic mutations without considering frequency in 59 PCC/PGL samples of TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma, *Del* deletion, *Ins* insertion

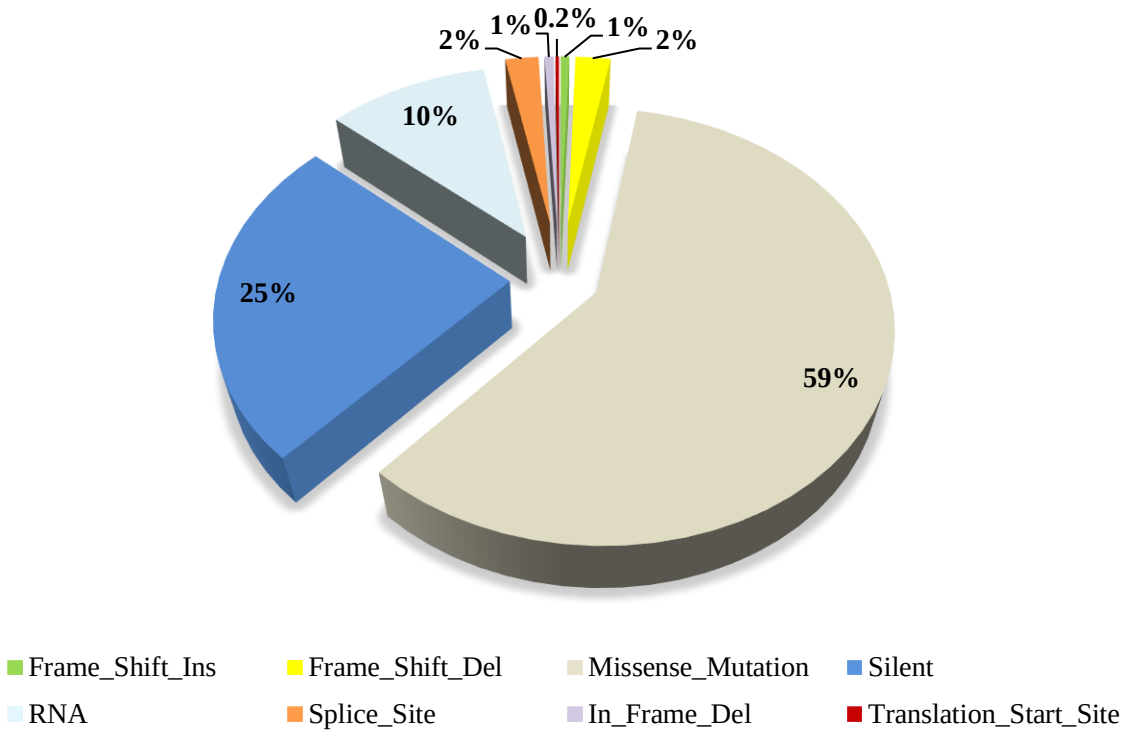


Table 7 Mutations significantly more frequent in malignant PCC/PGLs of TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma, *M* malignant, *B* benign

Genes	M w/ mutation (-)	M w/ mutation (+)	B w/ mutation (-)	B w/ mutation (+)	p value
ATRX	9	3	46	1	0.024
SETD2	10	2	47	0	0.039
MUC16	9	3	47	0	0.007
PAX1	10	2	47	0	0.039

Figure 7 Genetic alteration correlated with personal neoplasm status in 59 PCC/PGL samples of TCGA. *PCC* pheochromocytoma, *PGL* paraganglioma

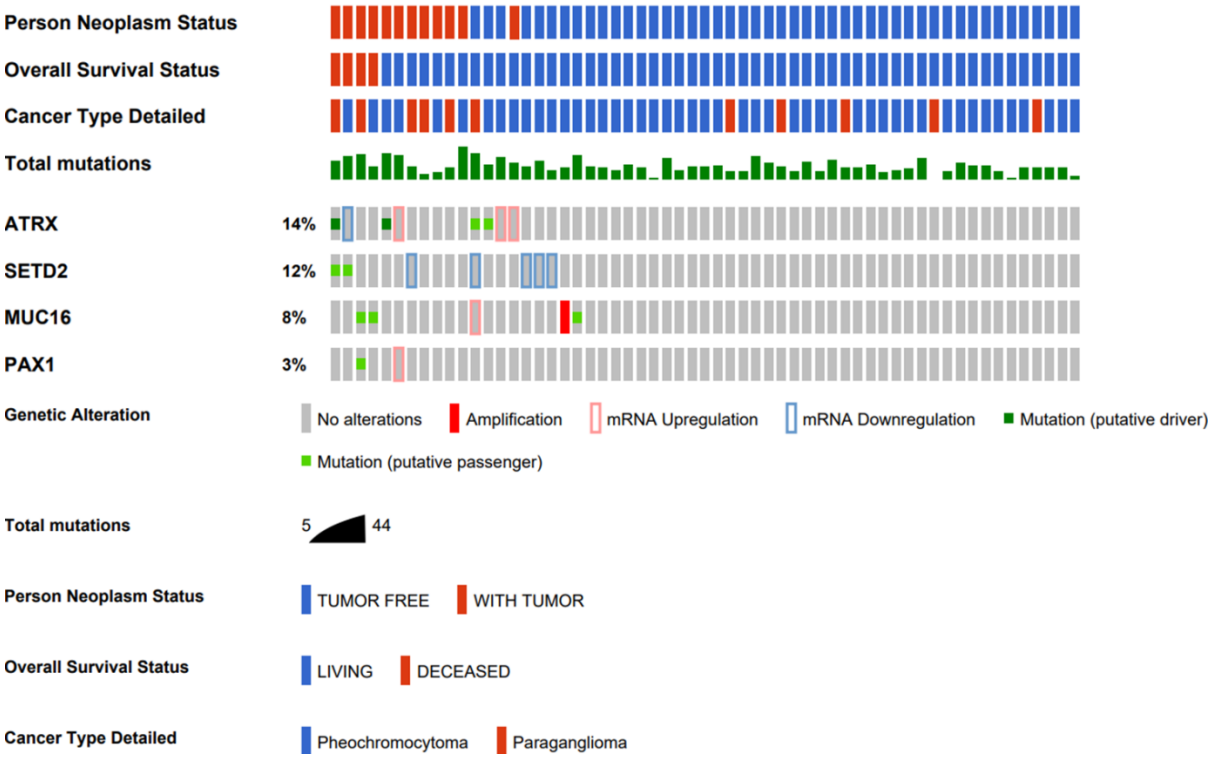


Figure 8 Kaplan–Meier curves for 59 patients from TCGA showing the effect of mutations on overall survival rates: (a) *ATRX*, (b) *SETD2*, (c) *MUC16*, and (d) *PAX1*.

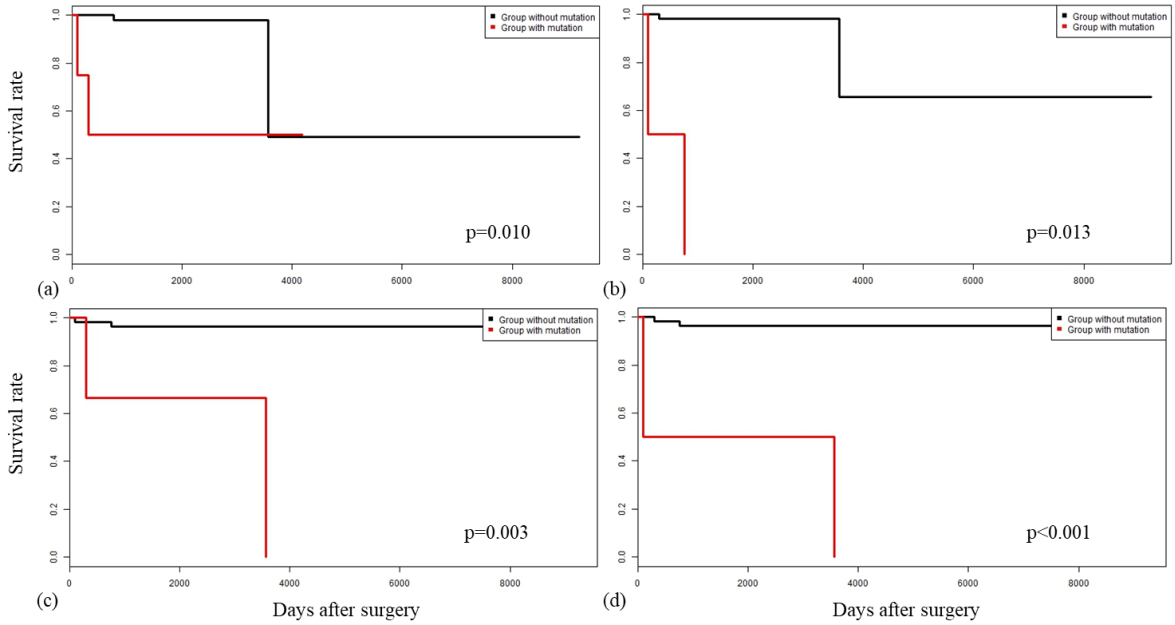


Figure 9 Kaplan-Meier curves for 59 patients from TCGA showing the most significant value when patients with at least one mutation of these mutations were compared to patients with none of these mutations ($p<0.001$).

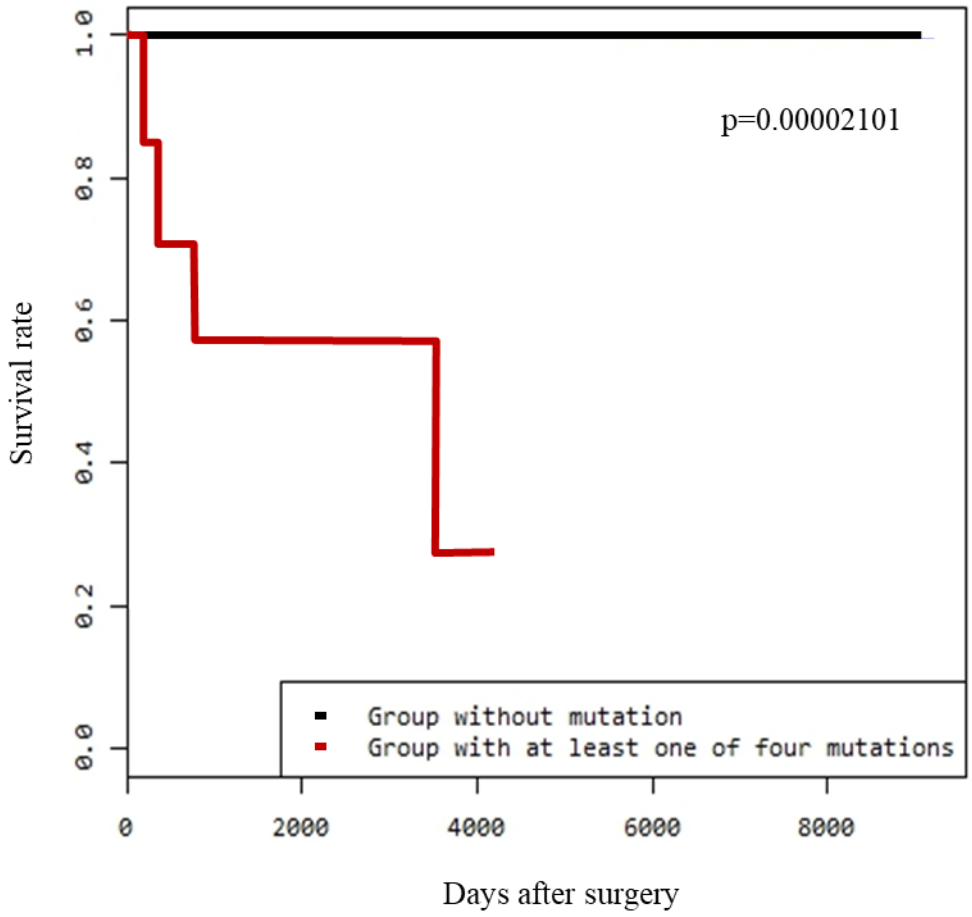
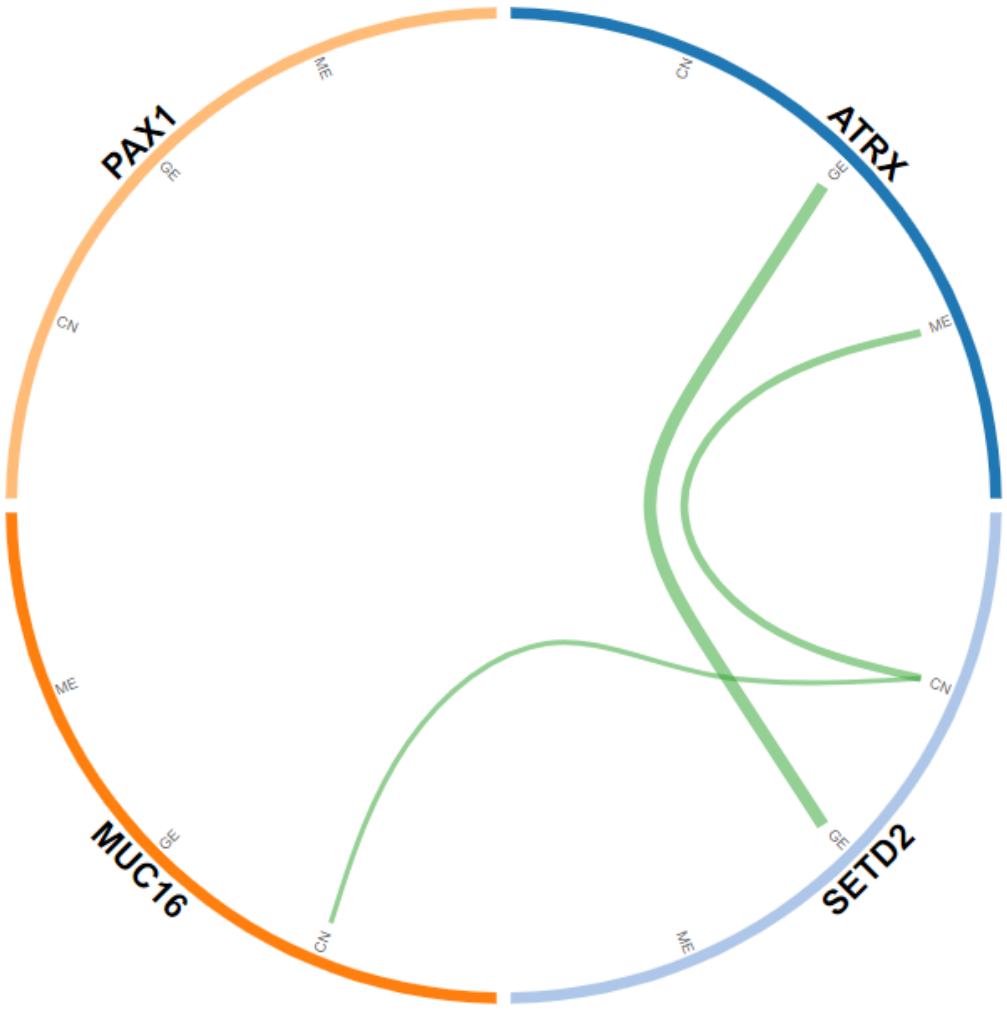


Figure 10 Interaction of 4 genes visualized in various platforms. Only significant intergenic interactions were shown (FDR ≤0.1). Green lines symbolize positive interactions. The thickness of green lines displays the strength of interactions. *GE* gene expression, *CN* copy number, *ME* methylation, *PE* protein expression



Although mutational analysis had limitations for data from EMBL-EBI, mutational rate of *ATRX* or *MUC16* was higher in the malignant group based on whole exome sequencing data of 31 PCC/PGLs pairs. However, data regarding *SETD2* or *PAX1* were unavailable. In targeted Sanger sequencing data of EMBL-EBI, no gene was found to be significantly frequent in malignant group. Analysis of all accessible data on germline mutations suggested significant malignant progression for *SDHB* mutation and potential malignancy for *VHL* mutation (Table 8).

Table 8 Comparison of germline mutations between malignant and benign PCC/PGLs from TCGA and EBML-EBI. *N/A* not available

Genes	TCGA		p value	EMBL-EBI		p value
	Malignant (n=11)	Benign (n=45)		Malignant (n=19)	Benign (n=165)	
EGLN1	0	1	1.000	N/A	N/A	—
MAX	0	0	1.000	0	1	1.000
NF1	0	2	1.000	0	12	0.617
RET	0	0	1.000	0	9	0.601
TMEM127	0	1	1.000	0	1	1.000
VHL	0	3	1.000	3	22	0.727
SDHB	5	3	0.0049	9	15	<0.001
SDHD	0	1	1.000	0	3	1.000
SDHA	N/A	N/A	—	0	1	1.000
SDHC	N/A	N/A	—	1	2	0.280

Discussion

Currently there are few reliable histopathologic criteria to predict malignant behavior in PCC/PGLs. Malignant potency and outcome varies with PCC/PGL genotypes. In clinical or pathologic aspects of the present study, possible risk factors included dopamine secretion, PGLs, and greater tumor size. These observations were in general agreement with previous studies [13, 14, 19, 27]. Genomic expression differences and mutational differences of malignant PCC/PGLs were investigated using TCGA and EMBL-EBI data for predicting prognosis in clinical practice. Thirty up/down-regulated pathways in malignant PCC/PGLs were associated with cancer signaling, metabolic alteration, prominent mitosis, and junctional dissociation. A total of 21 up-regulated genes and 11 down-regulated genes were significantly enriched in functional annotation pathways. In PCC/PGLs transformed to malignant types, cellular or nuclear proliferation was an indispensable process and down-regulation of cellular junctional pathway was essential for epithelial-mesenchymal transition. In addition, 4 candidate genes with mutations were proposed as genes susceptible to malignancy while germline mutations of *SDHB* or *VHL* meant possible malignant PCC/PGLs in the present study. *MUC16* or *PAX1* was not yet known in PCC/PGLs. *MUC16* (alias *CA125*) has been found on the apical surface of the ectodermal epithelia, playing a role as barrier [28]. *PAX1* is silenced by methylation. It might prohibit cell proliferation or division [29].

PASS and GAPP have been developed to assign malignant risk on the basis of histopathology [30-32]. These two risk stratification systems had several common features, including tumor necrosis, high cellularity, large nest, and vascular or capsular invasion. These recurrent themes might carry some implied validity. However, these systems are limited by their intrinsic problems or the absence of consistent validation [30, 32].

In one intuitive study, 58 pheochromocytoma samples have bene analyzed in order to separate malignant samples [33]. Thirteen samples were classified as malignant because they

showed lymph node or distant metastasis. Genome-wide expression was profiled and 10 genes among 109 DEGs were selected [33]. The current study had a similar focus to that study. However, the present study was designed following the awareness of challenging problems written in the 8th edition of the AJCC staging system. Malignant tumors may not be associated with local invasion or locoregional lymph node metastases. Therefore, the stricter definition of malignancy was applied for classifying patients with PCC/PGLs. Data from two public databases were used for cross-validation. Even in functionally enriched pathways, each cascade was investigated for common genes working in the same direction. Thirty functional pathways were presumably altered to malignancy. There were 21 up-regulated and 11 down-regulated genes. They could be used as part of gene set for potential classifier sooner or later.

Accumulated biological information helped us produce current results. These genes were cross-validated. In other studies, tumorigenesis of PCC/PGLs has been explained by alteration of the following three representative molecular pathways [18, 21, 23, 34, 35]: 1) pseudohypoxia and aberrant VEGF signaling; 2) kinase signaling such as PI3 kinase/AKT, RAS/RAF/ERK, and mTORC1/p70S6K; and 3) WNT signaling. In the present study, DEGs were enriched in 30 presumptive pathways that were closely interactive with the three known molecular pathways mentioned above.

It is well-known that malignant transformation is associated with genetic aberrations. In sporadic PCC/PGLs, up to 24% of PCC/PGLs have genetic predisposition [1, 3, 19, 36]. Many previous studies have investigated germline mutations in susceptible genes and identified germline genes such as *EPAS1, FH, MAX, MDH2, NF1, RET, SDHA, SDHAF2, SDHAF2, SDHB, SDHC, SDHD, TMEM127,* and *VHL* [10, 15, 21, 34, 37-39]. In the current study, somatic mutations from TCGA were processed. According to molecular pathways, percentage of germline mutation is different from that of somatic mutation in terms of their

contribution to PCC/PGL tumorigenesis [18]. However, somatic mutations can make equivalent contribution to PCC/PGL tumorigenesis [25, 40]. Burnichon et al. have identified somatic mutations in *VHL* and *RET* genes [25]. Somatic mutations of *NF1* and *MAX* have also been depicted by Castro-Vega et al. [24]. In a previous preview, genetic diversity has been introduced in aspects of germline/mosaic/somatic mutation by the timing of mutation acquisition from germ cell to adulthood [40]. Somatic mutation arises only in *ATRX*, *CSDE1*, *FGFR1*, *HRAS*, *IDH1*, *MAML3*, and *SETD2*. On the other hand, germline mutation occurs only in *EGLN1*, *FH*, *MDH2*, *SDHAF2*, *SDH (A, B C, D)*, and *TMEM127* [40]. The present study in close agreement with the preview, proposing 4 candidate mutations as genes susceptible to malignancy.

To the best of our knowledge, this study is the first integrative analysis of TCGA and EMBL-EBI for malignant PCC/PGLs defined in the 8th edition of the AJCC staging system. We identified potential pathways leading to malignant transformation of PCC/PGLs and mutations frequent in malignant PCC/PGLs. Specific genes in tumors with metastasis are candidate markers that might be useful for screening and monitoring which can help us decide the extent or radicality of surgery and additional therapies. Novel molecular therapeutics can be considered following results of the current study.

The present study has a few limitations. First, thorough analysis of microRNA, DNA methylation, copy-number variation, and protein expression should be followed to ascertain the complete signature of malignant PCC/PGLs. Second, the duration of follow-up was relatively short. A longer monitoring time might be preferable because the clinical course of PCC/PGLs is unpredictable. Third, as a usual limitation of big data analysis, the present study suffered from obstacles to search intriguing information beyond given data.

Conclusions

Data from TCGA and EMBL-EBI showed differences in mRNA expression and mutations between malignant and benign PCC/PGLs. Results from our analyses will facilitate appropriate diagnosis and treatment for malignant PCC/PGLs.

References

1. Lenders JW, Eisenhofer G, Mannelli M, Pacak K. Phaeochromocytoma. Lancet 366: 665-75, 2005
2. Corssmit EP, Romijn JA. Clinical management of paragangliomas. Eur J Endocrinol 171: R231-43, 2014
3. Kiernan CM, Solorzano CC. Pheochromocytoma and Paraganglioma: Diagnosis, Genetics, and Treatment. Surg Oncol Clin N Am 25: 119-38, 2016
4. Joynt KE, Moslehi JJ, Baughman KL. Paragangliomas: etiology, presentation, and management. Cardiol Rev 17: 159-64, 2009
5. Bravo EL, Tagle R. Pheochromocytoma: state-of-the-art and future prospects. Endocr Rev 24: 539-53, 2003
6. Lenders JW, Duh QY, Eisenhofer G, et al. Pheochromocytoma and paraganglioma: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 99: 1915-42, 2014
7. Chrisoulidou A, Kaltsas G, Ilias I, Grossman AB. The diagnosis and management of malignant phaeochromocytoma and paraganglioma. Endocr Relat Cancer 14: 569-85, 2007
8. Baudin E, Habra MA, Deschamps F, et al. Therapy of endocrine disease: treatment of malignant pheochromocytoma and paraganglioma. Eur J Endocrinol 171: R111-22, 2014
9. Adjalle R, Plouin PF, Pacak K, Lehnert H. Treatment of malignant pheochromocytoma. Horm Metab Res 41: 687-96, 2009
10. Tischler AS, Pacak K, Eisenhofer G. The adrenal medulla and extra-adrenal paraganglia: then and now. Endocr Pathol 25: 49-58, 2014
11. Elder EE, Elder G, Larsson C. Pheochromocytoma and functional paraganglioma

syndrome: no longer the 10% tumor. J Surg Oncol 89: 193-201, 2005

12. Andersen KF, Altaf R, Krarup-Hansen A, et al. Malignant pheochromocytomas and paragangliomas - the importance of a multidisciplinary approach. Cancer Treat Rev 37: 111-9, 2011

13. Thompson LD. Pheochromocytoma of the Adrenal gland Scaled Score (PASS) to separate benign from malignant neoplasms: a clinicopathologic and immunophenotypic study of 100 cases. Am J Surg Pathol 26: 551-66, 2002

14. Goffredo P, Sosa JA, Roman SA. Malignant pheochromocytoma and paraganglioma: a population level analysis of long-term survival over two decades. J Surg Oncol 107: 659-64, 2013

15. Gimm O, DeMicco C, Perren A, Giammarile F, Walz MK, Brunaud L. Malignant pheochromocytomas and paragangliomas: a diagnostic challenge. Langenbecks Arch Surg 397: 155-77, 2012

16. de Krijger RR, van Nederveen FH, Korpershoek E, Dinjens WN. New developments in the detection of the clinical behavior of pheochromocytomas and paragangliomas. Endocr Pathol 17: 137-41, 2006

17. Scholz T, Schulz C, Klose S, Lehnert H. Diagnostic management of benign and malignant pheochromocytoma. Exp Clin Endocrinol Diabetes 115: 155-9, 2007

18. Bjorklund P, Pacak K, Crona J: Precision medicine in pheochromocytoma and paraganglioma: current and future concepts. J Intern Med 280: 559-73, 2016

19. Feng F, Zhu Y, Wang X, et al. Predictive factors for malignant pheochromocytoma: analysis of 136 patients. J Urol 185: 1583-90, 2011

20. Mannelli M, Castellano M, Schiavi F, et al. Clinically guided genetic screening in a large cohort of italian patients with pheochromocytomas and/or functional or nonfunctional paragangliomas. J Clin Endocrinol Metab 94: 1541-7, 2009

21. Favier J, Amar L, Gimenez-Roqueplo AP. Paraganglioma and phaeochromocytoma: from genetics to personalized medicine. Nat Rev Endocrinol 11: 101-11, 2015

22. Gimenez-Roqueplo AP, Burnichon N, Amar L, Favier J, Jeunemaitre X, Plouin PF. Recent advances in the genetics of phaeochromocytoma and functional paraganglioma. Clin Exp Pharmacol Physiol 35: 376-9, 2008

23. Fishbein L, Leshchiner I, Walter V, et al. Comprehensive Molecular Characterization of Pheochromocytoma and Paraganglioma. Cancer Cell 31: 181-93, 2017

24. Castro-Vega LJ, Letouze E, Burnichon N, et al. Multi-omics analysis defines core genomic alterations in pheochromocytomas and paragangliomas. Nat Commun 6:6044, 2015

25. Burnichon N, Vescovo L, Amar L, et al. Integrative genomic analysis reveals somatic mutations in pheochromocytoma and paraganglioma. Hum Mol Genet 20: 3974-85, 2011

26. Zhu Y, Xu Y, Helseth DL, Jr., et al. Zodiac: A Comprehensive Depiction of Genetic Interactions in Cancer by Integrating TCGA Data. J Natl Cancer Inst 107: 8, 2015

27. Hamidi O, Young WF, Jr., Iniguez-Ariza NM, et al. Malignant Pheochromocytoma and Paraganglioma: 272 Patients Over 55 Years. J Clin Endocrinol Metab 102: 3296-305, 2017

28. Shah K, Patel S, Mirza S, Rawal RM. Unravelling the link between embryogenesis and cancer metastasis. Gene 642: 447-52, 2018

29. Juodzbalys G, Kasradze D, Cicciu M, et al. Modern molecular biomarkers of head and neck cancer. Part I. Epigenetic diagnostics and prognostics: Systematic review. Cancer Biomark 17: 487-502, 2016

30. Tischler AS, de Krijger RR. 15 YEARS OF PARAGANGLIOMA: Pathology of pheochromocytoma and paraganglioma. Endocr Relat Cancer 22: T123-33, 2015

31. Kimura N, Watanabe T, Noshiro T, Shizawa S, Miura Y. Histological grading of adrenal and extra-adrenal pheochromocytomas and relationship to prognosis: a clinicopathological analysis of 116 adrenal pheochromocytomas and 30 extra-adrenal sympathetic paragangliomas including 38 malignant tumors. Endocr Pathol 16: 23-32, 2005

32. Eisenhofer G, Tischler AS. Neuroendocrine cancer. Closing the GAPP on predicting metastases. Nat Rev Endocrinol 10: 315-6, 2014

33. Suh I, Shibru D, Eisenhofer G, et al. Candidate genes associated with malignant pheochromocytomas by genome-wide expression profiling. Ann Surg 250: 983-90, 2009

34. Nolting S, Grossman AB. Signaling pathways in pheochromocytomas and paragangliomas: prospects for future therapies. Endocr Pathol 23: 21-33, 2012

35. Santarpia L, Habra MA, Jimenez C. Malignant pheochromocytomas and paragangliomas: molecular signaling pathways and emerging therapies. Horm Metab Res 41: 680-6, 2009

36. Amar L, Bertherat J, Baudin E, et al. Genetic testing in pheochromocytoma or functional paraganglioma. J Clin Oncol 23: 8812-18, 2005

37. Jafri M, Maher ER. The genetics of phaeochromocytoma: using clinical features to guide genetic testing. Eur J Endocrinol 166: 151-58, 2012

38. Burnichon N, Buffet A, Gimenez-Roqueplo AP. Pheochromocytoma and paraganglioma: molecular testing and personalized medicine. Curr Opin Oncol 28: 5-10, 2016

39. Conzo G, Pasquali D, Colantuoni V, et al. Current concepts of pheochromocytoma. Int J Surg 12: 469-74, 2014

40. Dahia PL. Pheochromocytomas and Paragangliomas, Genetically Diverse and

Minimalist, All at Once! Cancer Cell 31:159-61, 2017

국문초록

서론: 갈색세포종과 부신경절종의 악성 여부를 예측할 방법이 필요하다. 하지만 악성 갈색세포종과 부신경절종을 가려낼 믿을 만한 조직병리학적 기준은 부족하다. 최근 공용 빅 데이터인 The Cancer Genome Atlas (TCGA)와 European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI)의 등장은 질병에 대해 보다 정확히 이해할 수 있는 유전체 정보를 제공한다. 악성 갈색세포종과 부신경절종의 유전체 분석에도 이 TCGA와 EMBL-EBI를 활용할 수 있다. 이 연구에서는 TCGA와 EMBL-EBI 데이터를 이용하여 악성 갈색세포종과 부신경절종의 mRNA 발현 및 돌연변이의 특징을 규명하고자 한다.

방법: TCGA와 EMBL-EBI는 각각 184개와 202개의 갈색세포종과 부신경절종 검체에 대해 유전체 분석 데이터를 가지고 있다. 이에 TCGA와 EBML-EBI의 임상 정보, 돌연변이, mRNA 발현 정보를 악성 갈색세포종과 부신경절종 분석에 활용하였다. 포함/제외 기준을 따라 TCGA에서는 59개의 검체가, EMBL-EBI에서는 171개의 검체가 연구에 포함되었다. 59개의 TCGA 검체 중에서 12개는 악성이었으며 47개는 양성이었다. 171개의 EMBL-EBI 검체 중에서는 19개가 악성이었으며 152개는 양성이었다. mRNA 발현과 돌연변이를 이 두 그룹에서 비교하였다.

결과: 악성 갈색세포종과 부신경절종은 30개의 세포 경로에서 의미있는 변화가 관찰되었으며, 이는 암 신호전달경로, 대사변화, 세포분열, 접합해리와 연관되어 있었다. 21개의 유전자가 상향 조절되었으며 11개의 유전자가 하향 조절되었다. *SDHB* 또는 *VHL*의 유전자에서 생식세포 돌연변이가 생길 경우 악성 관련성이 있었다. 또한, 악성 그룹은 *ATRX*, *SETD2*, *MUC16*, *PAX1*의 유전자에서 체세포

돌연변이가 빈번하게 나타났다. 이 유전자들의 돌연변이 유무는 전체 생존율과도 유의한 상관관계가 있었다.

결론: TCGA와 EMBL-EBI의 갈색세포종과 부신경절종 데이터는 악성 여부에 따른 mRNA 발현과 돌연변이의 차이를 보여주었다. 갈색세포종과 부신경절종의 유전체에 대한 발전된 이해는 올바른 진단과 적절한 치료를 제공하는 데 도움을 줄 수 있다.

주요어: 악성, 갈색세포종, 부신경절종, TCGA, EMBL-EBI

학 번: 2015-30562