Genetics and Molecular Research

E2F, HSF2, and miR-26 in thyroid carcinoma: bioinformatic analysis of RNA-sequencing data

J.C. Lu and Y.P. Zhang

Department of Endocrinology, Changhai Hospital, Shanghai, China

Corresponding author: Y.P. Zhang
E-mail: ypzhadgj@163.com

Genet. Mol. Res. 15 (1): gmr.15017576
Received September 2, 2015
Accepted December 4, 2015
Published March 11, 2016
DOI http://dx.doi.org/10.4238/gmr.15017576

ABSTRACT. In this study, we examined the molecular mechanism of thyroid carcinoma (THCA) using bioinformatics. RNA-sequencing data of THCA (N = 498) and normal thyroid tissue (N = 59) were downloaded from The Cancer Genome Atlas. Next, gene expression levels were calculated using the TCC package and differentially expressed genes (DEGs) were identified using the edgeR package. A co-expression network was constructed using the EBcoexpress package and visualized by Cytoscape, and functional and pathway enrichment of DEGs in the co-expression network was analyzed with DAVID and KOBAS 2.0. Moreover, modules in the co-expression network were identified and annotated using MCODE and BiNGO plugins. Small-molecule drugs were analyzed using the cMAP database, and miRNAs and transcription factors regulating DEGs were identified by WebGestalt. A total of 254 up-regulated and 59 down-regulated DEGs were identified between THCA samples and controls. DEGs enriched in biological process terms were related to cell adhesion, death, and growth and negatively correlated with various small-molecule drugs. The co-expression network of the DEGs consisted of hub genes (*ITGA3*, *TIMP1*, *KRT19*, and *SERPINA1*) and one module (*JUN*, *FOSB*, and *EGR1*). Furthermore, 5 miRNAs and 5 transcription factors were

identified, including E2F, HSF2, and miR-26. miR-26 may participate in THCA by targeting *CITED1* and *PLA2R1*; E2F may participate in THCA by regulating *ITGA3*, *TIMP1*, *KRT19*, *EGR1*, and *JUN*; HSF2 may be involved in THCA development by regulating *SERPINA1* and *FOSB*; and small-molecule drugs may have anti-THCA effects. Our results provide novel directions for mechanistic studies and drug design of THCA.

Key words: Differentially expressed genes; MicroRNAs; RNA sequencing; Thyroid carcinoma; Transcription factors

INTRODUCTION

Thyroid carcinoma (THCA) is a prevalent endocrine malignancy (Xing, 2013). Over the past several decades, the incidence of THCA has been steadily increased worldwide. Particularly, the THCA mortality rates in most countries ranged between 0.20-0.40 per 100,000 men and 0.20-0.60 per 100,000 women in 2008-2012 (La Vecchia et al., 2015). Currently, available treatment involves surgery, radioactive iodine therapy, and thyroid hormone. However, the median survival is reported to be 3-5 months from diagnosis, and the carcinoma worsens when chemotherapy fails (Sosa et al., 2014).

Currently, various genetic and epigenetic alterations have been found to participate in the initiation and progression of THCA, including activation of the phosphatidylinositol 3-kinase-protein kinase B (AKT) and mitogen-activated protein kinase signaling pathways, point mutations in *BRAF* and *RAS*, and chromosomal rearrangement of paired box 8/peroxisome proliferator-activated receptor γ and *RET*/*PTC* (Nikiforov and Nikiforova, 2011). As an effective gene expression regulator at the post-transcription level, microRNA (miRNA) has been found to play crucial roles in THCA; for example, 1) miR-145 is significantly down-regulated in THCA, and its overexpression inhibits the growth and metastasis of THCA by targeting *AKT3* and inhibiting the PI3K/Akt pathway (Boufraqech et al., 2014); 2) miR-146b-5p promotes metastasis and the epithelial-mesenchymal transition in papillary THCA by targeting zinc RING finger 3 and enhancing Wnt/β-catenin signaling (Deng et al., 2015); 3) down-regulation of miR-181b accelerates apoptosis by targeting *CYLD* in papillary THCA (Li et al., 2014).

In addition, transcription factors (TFs) regulate gene expression at the transcription level and participate in THCA; for example, 1) as a basic helix-loop-helix TF, Twist1 affects the survival and motility of THCA cells by up-regulating *HS6ST2, COL1A1, F2RL1, LEPREL1, PDZK1*, and *PDZK1IP1* (Di Maro et al., 2014) and 2) Runx2 promotes the invasion and metastasis of THCA by regulating epithelial-mesenchymal transition-related molecules, angiogenic/lymphangiogenic factors, and matrix metalloproteinases (Niu et al., 2012). However, the molecular mechanism underlying THCA is not well understood.

To gain insight into the genetic changes that occur in THCA, RNA sequencing (RNA-seq) data were downloaded from a public database for analysis. We conducted differential expression analysis, co-expression analysis, and functional annotation to detect all possible changes. To comprehensively understand these changes, differentially expressed genes (DEGs) related small-molecule drugs, miRNAs and TFs were screened. The results of this study may provide novel directions for mechanism studies and drug design for THCA.

MATERIAL AND METHODS

Preprocessing of RNA-seq data

All second generation RNA-seq data of THCA from The Cancer Genome Atlas (TCGA) were analyzed using the TCGA-Assembler software (Zhu et al., 2014), and only solid tumor (THCA group, N = 498) and normal solid tissue (control group, N = 59) samples were included in this study. Next, the reads per kilobase per million mapped sequence read values, which could be used to evaluate gene expression levels, were calculated and normalized using the Tag count comparison package (version1.6.5, http://www.bioconductor.org/packages/release/bioc/html/TCC.html) (Sun et al., 2013).

DEG screening

DEGs between THCA and control samples were identified using the edgeR package from Bioconductor (version3.8.6, http://www.bioconductor.org/packages/3.0/bioc/html/edgeR.html) (Robinson et al., 2010). In this analysis, the gene expression values of THCA and control samples were evaluated based on edgeR models, and P values were adjusted using the multtest package (version 2.22.0, http://www.bioconductor.org/packages/3.0/bioc/html/multtest.html) (Benjamini, 2010) to obtain the false discovery rate (FDR). The criteria for DEG screening were FDR < 0.05 and $|\log_2$-fold change (FC)|>1, in which FC = gene expression value in THCA group / gene expression value in the control group.

Bidirectional hierarchical clustering analysis of DEGs

In bidirectional hierarchical clustering analysis (Boufraqech et al., 2014), the gene symbols and expression values of DEGs in every sample were extracted and clustered using the pheatmap package (version1.0.2, http://cran.r-project.org/web/packages/pheatmap/index.html) (Kolde, 2012) based on Euclidean distance (Deza and Deza, 2009). Next, the results of this analysis were illustrated using heatmap.

Construction of co-expression network

The correlations between DEGs were calculated using the EBcoexpress package from Bioconductor (version1.10.0, http://www.bioconductor.org/packages/3.0/bioc/html/EBcoexpress.html) (Dawson et al., 2012), and the DEG-DEG pairs with correlation coefficient (|r|) > 0.6 were utilized to construct the co-expression network, which was visualized using Cytoscape (version 2.8, http://cytoscape.org/) (Smoot et al., 2011).

Functional and pathway enrichment analyses of DEGs

Gene annotation information as downloaded from the Gene Ontology (GO) Consortium (http://geneontology.org/) and DEGs were annotated and given GO identifiers, which were further divided into Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories (Blake et al., 2013). GO functional enrichment of DEGs in the co-expression network was also conducted, accompanied by Kyoto Encyclopedia of Genes and Genomes pathway en-

richment using the online tool DAVID (http://david.abcc.ncifcrf.gov/) (Huang et al., 2007) and web server KOBAS 2.0 (http://kobas.cbi.pku.edu.cn) (Xie et al., 2011) based on a hypergeometric algorithm. The threshold for these analyses was set as P value < 0.05.

Functional annotation of modules in co-expression network

The modules in the co-expression network were identified and annotated using the molecular complex detection plugin (Bader and Hogue, 2003) and biological networks gene ontology tool plugin (Maere et al., 2005) in Cytoscape (http://www.cytoscape.org/) (Smoot et al., 2011), respectively. The criteria for module division were degree cutoff ≥ 2 (the connective degree of every node in the module should ≥ 2) and K-core ≥ 2 (the neighbor nodes of every node in the module should ≥ 2). Annotation was performed based on the hypergeometric distribution with a threshold of adjusted P value < 0.01.

Small-molecule drug analysis of DEGs

In order to identify functional relationships between THCA, DEGs, and bioactive small-molecule drugs, small-molecule drug analysis was performed for the identified DEGs based on the Connectivity map database (http://www.broadinstitute.org/cmap/) (Lamb et al., 2006). The Connectivity map database stores genome-wide expression data from cultured human cells treated with bioactive small molecules. The criterion for this analysis was correlation coefficient (|score|) > 0.8.

miRNAs and TFs analysis of DEGs

The miRNAs and TFs regulating DEGs were identified using online software WEB-based GEne SeT AnaLysis Toolkit (http://bioinfo.vanderbilt.edu/webgestalt/) (Wang et al., 2013), and the corresponding criterion was set as adjusted P value < 0.05 via multiple test.

RESULTS

DEGs between THCA and control groups

A total of 11,483 genes were detected in the RNA-seq data of THCA samples and controls. After DEG screening, a total of 313 significant DEGs (FDR < 0.05 and $|\log_2 FC|>1$) were identified between THCA and control groups, including 254 up-regulated and 59 down-regulated DEGs. Particularly, *CITED1* ($\log_2 FC = 4.97$, FDR = 0) and*PLA2R1* ($\log_2 FC = -2.93$, FDR = 0) were among the most significantly down-regulated DEGs.

Bidirectional hierarchical clustering analysis of DEGs

After bidirectional hierarchical clustering analysis, DEGs between 498 THCA samples and 59 controls were clustered based on their expression values, and then illustrated using heatmap (Figure 1). THCA samples and controls could be clearly differentiated using the identified DEGs.

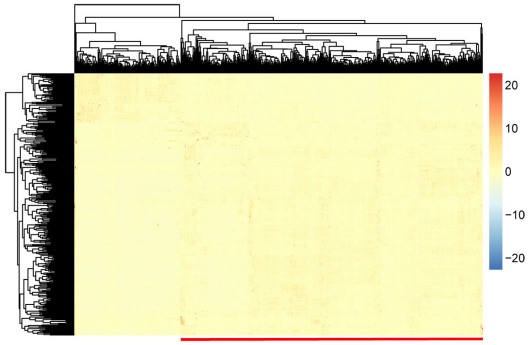

Figure 1. Heatmap of DEGs between THCA and control groups. The red line under the heatmap indicates the THCA samples. The color bar on the right side represents gene expression values. DEGs: differentially expressed genes; THCA: thyroid carcinoma.

Functional annotation of DEGs

After GO functional annotation, 313 DEGs were predicted to participate in 11 CC terms, 12 MF terms, and 23 BP terms (Figure 2). The 11 CC terms were mainly associated with extracellular region, membrane, and synapse; the 12 MF terms were mainly related with gene expression, reaction activity, and transporter; and the 23 BP terms were mainly associated with cell adhesion, death, and growth. The terms chemoattractant activity, electron carrier activity, translation regulator activity, and cell killing only included up-regulated DEGs.

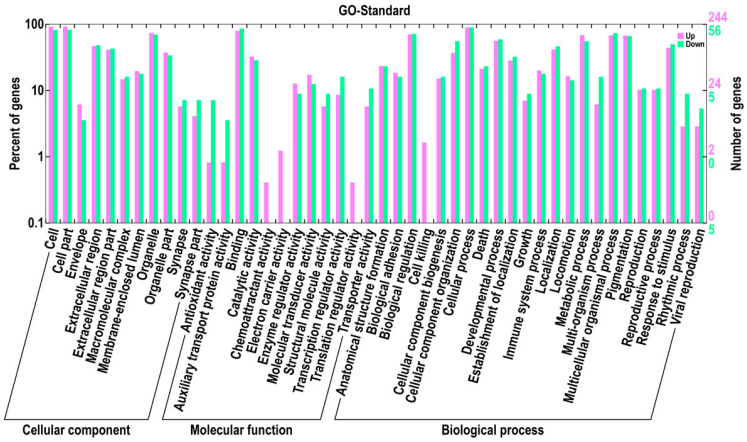

Figure 2. Results of functional annotation of DEGs. Pink: up-regulated DEGs; green: down-regulated DEGs; left vertical axis: percent of DEGs involved in a specific term against all DEGs; right vertical axis: number of DEGs involved in a specific term; DEGs: differentially expressed genes.

Co-expression network of DEGs

Based on the DEG-DEG pairs with |r| > 0.6, the co-expression network of DEGs was constructed, consisting of 20 DEGs (17 up-regulated DEGs and 3 down-regulated DEGs) and 21 co-expression relationships (Figure 3). In this network, the up-regulated DEGs *ITGA3* (degree =

6, $\log_2$ FC = 1.70), *TIMP1* (degree = 3, $\log_2$ FC = 3.05), *KRT19* (degree = 3, $\log_2$ FC = 2.73), and *SERPINA1* (degree = 3, $\log_2$ FC = 4.61) had a high connective degree, and thus were defined as hub genes.

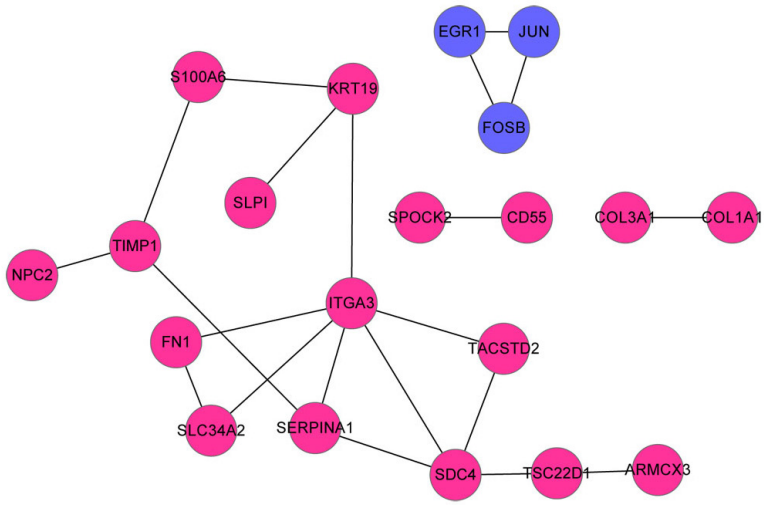

Figure 3. Co-expression network of DEGs. Red nodes: up-regulated DEGs; blue nodes: down-regulated DEGs; edges: co-expression relationship between DEGs; DEGs: differentially expressed genes.

Enrichment analysis of DEGs in co-expression network

After GO functional and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses, DEGs in the co-expression network were significantly enriched in 10 functions, mainly including response to hormone stimulus, adhesion, and collagen biosynthesis (Figure 4) and 2 pathways, including extracellular matrix-receptor interaction and focal adhesion (P value < 0.05).

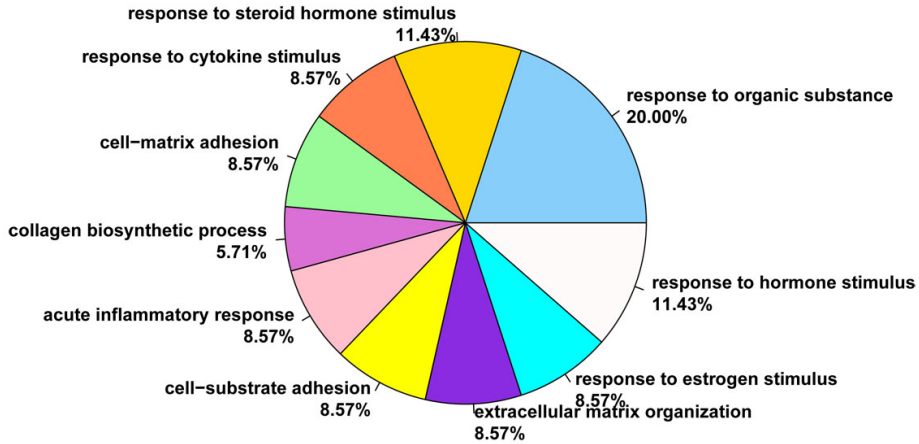

Figure 4. GO functions involving DEGs in co-expression network. The percent in the figure represents the percent of DEGs in a specific term against all DEGs. GO: Gene Ontology; DEGs: differentially expressed genes.

Functional annotation of modules in co-expression network

In the co-expression network, only one module was identified (degree cutoff ≥ 2 and K-core ≥ 2), consisting of the down-regulated DEGs *JUN* (log_2 FC = -1.73), *FOSB* (log_2 FC = -2.30), and *EGR1* (log_2 FC = -1.72), and 3 co-expression relationships. Furthermore, the DEGs in this module were annotated to participate in 8 GO functions, which were mainly associated with transcription (adjusted P value < 0.01, Table 1).

Table 1. Functional annotation of DEGs in the identified module.

GOID	Adjusted P value	DEG	Gene function
7610	8.22E-03	JUN, FOSB, EGR1	Behavior
6357	8.22E-03	JUN, FOSB, EGR1	Regulation of transcription from RNA polymerase II promoter
31324	8.22E-03	JUN, FOSB, EGR1	Negative regulation of cellular metabolic process
10605	8.22E-03	JUN, FOSB, EGR1	Negative regulation of macromolecule metabolic process
9892	8.22E-03	JUN, FOSB, EGR1	Negative regulation of metabolic process
6355	1.97E-02	JUN, FOSB, EGR1	Regulation of transcription, DNA-dependent
48523	1.97E-02	JUN, FOSB, EGR1	Negative regulation of cellular process
51252	1.97E-02	JUN, FOSB, EGR1	Regulation of RNA metabolic process

GO: Gene Ontology; ID: identifier; DEGs: differentially expressed genes.

Small-molecule drug analysis of DEGs

After small-molecule drug analysis, the DEGs between the THCA and control groups were significantly correlated with 11 small-molecule drugs (|score| > 0.8, Table 2). Eight drugs were negatively correlated with the identified DEGs, including doxylamine, Prestwick-920, cromoglicic acid, nimesulide, and pridinol.

Table 2. Small-molecule drugs correlated with DEGs.

Small-molecule drugs	Correlation score	P value
Cobalt chloride	-0.882	0.00318
Doxylamine	-0.879	0.00008
Prestwick-920	-0.877	0.00052
Cromoglicic acid	-0.861	0.03887
Nimesulide	-0.844	0.00105
AG-012559	-0.824	0.01098
AH-23848	-0.821	0.01146
Pridinol	-0.817	0.00209
Cefamandole	0.836	0.00111
Amantadine	0.85	0.00072
Fenofibrate	0.864	0.00465

The correlation score changes from 1 to -1. Score < 0 means the change directions of gene expression values caused by this drug are opposite to the change directions of gene expres-

sion values in THCA, namely, this drug might have anti-THCA effects. What's more, low scores represent high anti-THCA effects. On the contrary, score > 0 means the change directions of gene expression values caused by this drug are same to the change directions of gene expression values in THCA, namely, this drug might promote THCA. In addition, high scores represent high THCA-promoting effects. THCA: thyroid carcinoma; DEGs: differentially expressed genes.

miRNAs and TFs regulating DEGs

A total of 5 miRNAs and 5 TFs were found to regulate DEGs (adjusted P value < 0.05, Table 3). Particularly, 1) E2F was found to regulate hub genes (*ITGA3*, *KRT19*, *TIMP1*) and module genes (*EGR1*, *JUN*); 2) HSF2 was found to regulate the hub gene *SERPINA1* and the module gene *FOSB*; 3) miR-26 was found to regulate the most significant DEGs *CITED1* and *PLA2R1*.

Table 3. miRNAs and TFs regulating DEGs.

Regulator	Regulator name	Regulator ID	Adjusted P value	DEGs
TF	hsa-E2F	DB_ID:2297	3.15E-05	*S100A11, NPC2, TGM2, LGALS1, CYR61, ENC1, SOX4, JUN, ITGA3, CTSC, TGFBI, ELF3, FHL1, TNC, KRT19, TIMP1, EGR1, PNP, ECE1, FBLN1, EFEMP1, COL3A1, DCN*
	hsa-HSF2	DB_ID:1951	5.98E-04	*CTSH, COL1A1, MATN2, SPOCK1, CYP1B1, ERBB3, PRSS23, DHRS3, SPOCK2, GALE, CRABP2, EPS8, ICAM1, P4HA2, FOSB, E2F1, SPINT1, SERPINA1*
	hsa-VDR	DB_ID:2102	2.63E-04	*SLPI, PAPSS2, ALDH1A3, S100A4, UPP1, CDH3, LAD1, COL5A1, ETV5, PSD3, APOE, TBC1D4, NELL2, CDH2, MET*
	hsa-HNF3	DB_ID:2423	3.11E-04	*CHST2, CA12, DUSP4, NRCAM, SLC34A2, S100A2, SORBS2, TMC6, APOC1, ALOX5, PC, BID, ELMO1, VCAN, NFE2L3, CDH13, CSPG4, PCSK2, TLE4, GJA4*
	hsa-AP4	DB_ID:1967	1.00E-04	*MTMR11, MAMLD1, QPCT, IL1RAP, EGR2, INHBB, BMP2, PHYHIP, S100A1, PLAG1, CFD, MST1R, PLAU, ADORA1*
miRNA	hsa-miR-520	DB_ID:699	2.00E-04	*PLAG1, CFD, MST1R, PLAU, ADORA1, RYR1, SCN1B, GLS2, CDH6, CORO2A, HRH1, PDE9A, ALDH3B1, ODZ1, PRSS22*
	hsa-miR-524	DB_ID:802	4.50E-03	*KL, EPHA3, TPPP, TNFRSF10C, TIAM1, SLC26A4, PTP4A3, PASK, CST6, PDE5A, KCNJ2*
	hsa-miR-26	DB_ID:687	2.50E-03	*SCEL, TGFBR1, CITED1, PLA2R1, CDKN2B, ENTPD1, IGSF1, PROS1, MUC1, BMP8A, CITED2, MGAT3*
	hsa-miR-124A	DB_ID:811	2.50E-03	*ESM1, CCND1, DUSP6, SLC1A5, LGALS3, MDK, HBB, AGR2*
	hsa-miR-15	DB_ID:811	5.00E-03	*ECM1, TM4SF1, CHI3L1, SYNE1, DUSP5, RAB27A*

DEGs: differentially expressed genes; miRNAs: microRNAs; TFs: transcription factors; ID: identifier.

DISCUSSION

THCA is an endocrine malignant cancer with a high incidence worldwide (Xing, 2013; La Vecchia et al., 2015). In this study, RNA-seq data were downloaded from public databases and reanalyzed comprehensively. A total of 254 up-regulated and 59 down-regulated DEGs were identified (*CITED1*, *PLA2R1*, etc.) and were negatively correlated with doxylamine, Prestwick-920, cromoglicic acid, nimesulide, and pridinol. A co-expression network of DEGs was constructed, which included hub genes (*ITGA3*, *TIMP1*, *KRT19*, and *SERPINA1*); next, one module (*JUN*, *FOSB*, and *EGR1*) was further evaluated. Furthermore, 5 miRNAs and 5 TFs were identified, including E2F (*ITGA3, KRT19, TIMP1, EGR1,* and *JUN*), HSF2 (*SERPINA1* and *FOSB*), and miR-26 (*CITED1* and *PLA2R1*), as regulators of DEGs.

Among the most significant DEGs identified, *CITED1* (Cbp/p300-interacting transactivator

1) has been reported to be greatly up-regulated in papillary THCA compared to in the normal thyroid (Prasad et al., 2004), and it plays diagnostic role in encapsulated lesions of papillary THCA (Scognamiglio et al., 2006). PLA2R1 (phospholipase A2 receptor 1) is down-regulated in malignant follicular THCA compared with in benign follicular adenomas (Fryknäs et al., 2006). Both *CITED1* and *PLA2R1* were regulated by miR-26. Decreased expression of miR-26 was previously detected in anaplastic THCA (Braun et al., 2010). Thus, miR-26 may participate in THCA by targeting *CITED1* and *PLA2R1*.

Among the hub genes and module genes, *ITGA3* (integrin α3), *TIMP1* (TIMP metallopeptidase inhibitor 1), *KRT19* (keratin 19), *EGR1* (early growth response 1), and *JUN* (Jun proto-oncogene) were regulated by E2F. *TIMP1* was found to be up-regulated in THCA, and the protein product binds to its receptor on cell membrane surface to activate the Akt signaling pathway, which confers anti-apoptotic behavior and promotes cell invasion (Bommarito et al., 2011). *KRT19* is overexpressed in papillary THCA (Cheung et al., 2001), and our result is consistent with those of a previous study. *EGR1* is expressed in the nucleus and up-regulates PTEN, a tumor suppressor that inhibits the cell cycle (Di Loreto et al., 2005). In this study, *EGR1* expression was down-regulated in THCA, promoting tumor growth. Additionally, the TF E2F plays crucial roles in regulating the G1/S transition during the cell cycle by activating genes involved in DNA synthesis and is up-regulated in benign and malignant thyroid tumors compared with in normal thyroid tissue (Saiz et al., 2002). Associations between E2F and *TIMP1*, *KRT19*, and *EGR1* have also been reported (Bigelow et al., 2009). Differential expression of *ITGA3* and *JUN* was identified in cancer tissues (Briggs et al., 2002; Bredel et al., 2005), and the expression differed between THCA samples and controls. *ITGA3* and *JUN* were enriched in the focal adhesion pathway, which is crucial for THCA invasion. Therefore, we predicted that *ITGA3* and *JUN* play a role in THCA development, and E2F participates in THCA by regulating *ITGA3*, *TIMP1*, *KRT19*, *EGR1*, and *JUN*.

Moreover, SERPINA1 (serpin peptidase inhibitor clade A member 1) and FOSB (FBJ murine osteosarcoma viral oncogene B) are regulated by HSF2 (heat shock transcription factor 2). Reportedly, *SERPINA1* is a biomarker of THCA (Griffith et al., 2006), and *FOSB* participates in known oncogenic pathways and various types of cancer (Milde-Langosch et al., 2003). Although there is little direct evidence regarding the involvement of HSF2 in THCA development, HSF2 has potential tumorigenic functions by regulating the proto-oncogene c-Fos (De Thonel et al., 2011). Therefore, HSF2 may participate in THCA development by regulating *SERPINA1* and *FOSB*.

Furthermore, the small-molecule drugs doxylamine, Prestwick-920, cromoglicic acid, nimesulide, and pridinol were significantly negatively correlated with the identified DEGs. Thus, the changes in the gene expression values caused by these drugs were opposite of the change of gene expression values in THCA, indicating that these drugs have anti-THCA effects.

In conclusion, 1) miR-26 may participate in THCA by targeting *CITED1* and *PLA2R1*; 2) *ITGA3* and *JUN* may play a role in THCA development; 3) E2F participates in THCA by regulating *ITGA3*, *TIMP1*, *KRT19*, *EGR1*, and *JUN*; 4) HSF2 may be involved in THCA development by regulating *SERPINA1* and *FOSB*; and 5) the small-molecule drugs doxylamine, Prestwick-920, cromoglicic acid, nimesulide, and pridinol may have anti-THCA effects. The results of this study will be useful for mechanism studies and drug design in THCA, and our further studies will focus on validating these predictions through experiments.

Conflicts of interest

The authors declare no conflict of interest.

REFERENCES

Bader GD and Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4: 2. http://dx.doi.org/10.1186/1471-2105-4-2

Benjamini Y (2010). Discovering the false discovery rate. *J. R. Stat. Soc. Series B Stat. Methodol.* 72: 405-416. http://dx.doi.org/10.1111/j.1467-9868.2010.00746.x

Bigelow RL, Williams BJ, Carroll JL, Daves LK, et al. (2009). TIMP-1 overexpression promotes tumorigenesis of MDA-MB-231 breast cancer cells and alters expression of a subset of cancer promoting genes in vivo distinct from those observed *in vitro*. *Breast Cancer Res. Treat.* 117: 31-44. http://dx.doi.org/10.1007/s10549-008-0170-7

Blake JA, Dolan M, Drabkin H, Hill DP, et al.; Gene Ontology Consortium (2013). Gene Ontology annotations and resources. *Nucleic Acids Res.* 41: D530-D535. http://dx.doi.org/10.1093/nar/gks1050

Bommarito A, Richiusa P, Carissimi E, Pizzolanti G, et al. (2011). BRAFV600E mutation, TIMP-1 upregulation, and NF-κB activation: closing the loop on the papillary thyroid cancer trilogy. *Endocr. Relat. Cancer* 18: 669-685. http://dx.doi.org/10.1530/ERC-11-0076

Boufraqech M, Zhang L, Jain M, Patel D, et al. (2014). miR-145 suppresses thyroid cancer growth and metastasis and targets AKT3. *Endocr. Relat. Cancer* 21: 517-531. http://dx.doi.org/10.1530/ERC-14-0077

Braun J, Hoang-Vu C, Dralle H and Hüttelmaier S (2010). Downregulation of microRNAs directs the EMT and invasive potential of anaplastic thyroid carcinomas. *Oncogene* 29: 4237-4244. http://dx.doi.org/10.1038/onc.2010.169

Bredel M, Bredel C, Juric D, Harsh GR, et al. (2005). Functional network analysis reveals extended gliomagenesis pathway maps and three novel MYC-interacting genes in human gliomas. *Cancer Res.* 65: 8679-8689. http://dx.doi.org/10.1158/0008-5472.CAN-05-1204

Briggs J, Chamboredon S, Castellazzi M, Kerry JA, et al. (2002). Transcriptional upregulation of SPARC, in response to c-Jun overexpression, contributes to increased motility and invasion of MCF7 breast cancer cells. *Oncogene* 21: 7077-7091. http://dx.doi.org/10.1038/sj.onc.1205857

Cheung CC, Ezzat S, Freeman JL, Rosen IB, et al. (2001). Immunohistochemical diagnosis of papillary thyroid carcinoma. *Mod. Pathol.* 14: 338-342. http://dx.doi.org/10.1038/modpathol.3880312

Dawson JA, Ye S and Kendziorski C (2012). R/EBcoexpress: an empirical Bayesian framework for discovering differential co-expression. *Bioinformatics* 28: 1939-1940. http://dx.doi.org/10.1093/bioinformatics/bts268

De Thonel A, Mezger V and Garrido C (2011). Implication of heat shock factors in tumorigenesis: therapeutical potential. *Cancers (Basel)* 3: 1158-1181. http://dx.doi.org/10.3390/cancers3011158

Deng X, Wu B, Xiao K, Kang J, et al. (2015). MiR-146b-5p promotes metastasis and induces epithelial-mesenchymal transition in thyroid cancer by targeting ZNRF3. *Cell. Physiol. Biochem.* 35: 71-82. http://dx.doi.org/10.1159/000369676

Deza M and Deza E (2009). Encyclopedia of Distances. In: Encyclopedia of distances. Springer, Berlin, 94.

Di Loreto C, Tell G, Pestrin M, Pandolfi M, et al. (2005). PTEN and Egr-1 expression in thyroid proliferative lesions. *Cancer Lett.* 224: 105-109. http://dx.doi.org/10.1016/j.canlet.2004.10.014

Di Maro G, Orlandella FM, Bencivenga TC, Salerno P, et al. (2014). Identification of targets of Twist1 transcription factor in thyroid cancer cells. *J. Clin. Endocrinol. Metab.* 99: E1617-E1626. http://dx.doi.org/10.1210/jc.2013-3799

Fryknäs M, Wickenberg-Bolin U, Göransson H, Gustafsson MG, et al. (2006). Molecular markers for discrimination of benign and malignant follicular thyroid tumors. *Tumour Biol.* 27: 211-220. http://dx.doi.org/10.1159/000093056

Griffith OL, Melck A, Jones SJ and Wiseman SM (2006). Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. *J. Clin. Oncol.* 24: 5043-5051. http://dx.doi.org/10.1200/JCO.2006.06.7330

Huang DW, Sherman BT, Tan Q, Collins JR, et al. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* 8: R183. http://dx.doi.org/10.1186/gb-2007-8-9-r183

Kolde R (2012). Pheatmap: pretty heatmaps. *R package version 0.7*. 4.

La Vecchia C, Malvezzi M, Bosetti C, Garavello W, et al. (2015). Thyroid cancer mortality and incidence: a global overview. *Int. J. Cancer* 136: 2187-2195. http://dx.doi.org/10.1002/ijc.29251

Lamb J, Crawford ED, Peck D, Modell JW, et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313: 1929-1935. http://dx.doi.org/10.1126/science.1132939

Li D, Jian W, Wei C, Song H, et al. (2014). Down-regulation of miR-181b promotes apoptosis by targeting CYLD in thyroid papillary cancer. *Int. J. Clin. Exp. Pathol.* 7: 7672-7680.

Maere S, Heymans K and Kuiper M (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21: 3448-3449. http://dx.doi.org/10.1093/bioinformatics/bti551

Milde-Langosch K, Kappes H, Riethdorf S, Löning T, et al. (2003). FosB is highly expressed in normal mammary epithelia, but down-regulated in poorly differentiated breast carcinomas. *Breast Cancer Res. Treat.* 77: 265-275. http://dx.doi.org/10.1023/A:1021887100216

Nikiforov YE and Nikiforova MN (2011). Molecular genetics and diagnosis of thyroid cancer. *Nat. Rev. Endocrinol.* 7: 569-580. http://dx.doi.org/10.1038/nrendo.2011.142

Niu DF, Kondo T, Nakazawa T, Oishi N, et al. (2012). Transcription factor Runx2 is a regulator of epithelial-mesenchymal transition and invasion in thyroid carcinomas. *Lab. Invest.* 92: 1181-1190. http://dx.doi.org/10.1038/labinvest.2012.84

Prasad ML, Pellegata NS, Kloos RT, Barbacioru C, et al. (2004). CITED1 protein expression suggests Papillary Thyroid Carcinoma in high throughput tissue microarray-based study. *Thyroid* 14: 169-175. http://dx.doi.org/10.1089/105072504773297830

Robinson MD, McCarthy DJ and Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-140. http://dx.doi.org/10.1093/bioinformatics/btp616

Saiz AD, Olvera M, Rezk S, Florentine BA, et al. (2002). Immunohistochemical expression of cyclin D1, E2F-1, and Ki-67 in benign and malignant thyroid lesions. *J. Pathol.* 198: 157-162. http://dx.doi.org/10.1002/path.1185

Scognamiglio T, Hyjek E, Kao J and Chen YT (2006). Diagnostic usefulness of HBME1, galectin-3, CK19, and CITED1 and evaluation of their expression in encapsulated lesions with questionable features of papillary thyroid carcinoma. *Am. J. Clin. Pathol.* 126: 700-708. http://dx.doi.org/10.1309/044V86JN2W3CN5YB

Smoot ME, Ono K, Ruscheinski J, Wang PL, et al. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27: 431-432. http://dx.doi.org/10.1093/bioinformatics/btq675

Sosa JA, Elisei R, Jarzab B, Balkissoon J, et al. (2014). Randomized safety and efficacy study of fosbretabulin with paclitaxel/carboplatin against anaplastic thyroid carcinoma. *Thyroid* 24: 232-240. http://dx.doi.org/10.1089/thy.2013.0078

Souza CF (2010). Validação do envolvimento dos genes KRT6A, KRT19, MSLN e KLK8 por RT-PCR quantitativa em tempo real em carcinomas epidermóides de cabeça e pescoço. Master's thesis. Universidade de São Paulo, São Paulo, Brazil.

Sun J, Nishiyama T, Shimizu K and Kadota K (2013). TCC: an R package for comparing tag count data with robust normalization strategies. *BMC Bioinformatics* 14: 219. http://dx.doi.org/10.1186/1471-2105-14-219

Wang J, Duncan D, Shi Z and Zhang B (2013). WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* 41: W77-83. http://dx.doi.org/10.1093/nar/gkt439

Xie C, Mao X, Huang J, Ding Y, et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39: W316-22. http://dx.doi.org/10.1093/nar/gkr483

Xing M (2013). Molecular pathogenesis and mechanisms of thyroid cancer. *Nat. Rev. Cancer* 13: 184-199. http://dx.doi.org/10.1038/nrc3431

Zheng C, Ren Z, Wang H, Zhang W, et al. (2009). E2F1 Induces tumor cell survival via nuclear factor-kappaB-dependent induction of EGR1 transcription in prostate cancer cells. *Cancer Res.* 69: 2324-2331. http://dx.doi.org/10.1158/0008-5472.CAN-08-4113

Zhu Y, Qiu P and Ji Y (2014). TCGA-assembler: open-source software for retrieving and processing TCGA data. *Nat. Methods* 11: 599-600. http://dx.doi.org/10.1038/nmeth.2956