

## Genes and pathways identified in thyroid carcinoma based on bioinformatics analysis

J. W. YU<sup>1</sup>, W. MAI<sup>2</sup>, Y. L. CUI<sup>3</sup>, L. Y. KONG<sup>1,\*</sup>

<sup>1</sup>Department of Head and Neck Surgery, The Affiliated Tumor Hospital of Harbin Medical University, Harbin, 150081 China; <sup>2</sup>Department of Orthopedics, The Affiliated Tumor Hospital of Harbin Medical University, Harbin, 150081 China; <sup>3</sup>Department of Radiation Oncology, The Affiliated Tumor Hospital of Harbin Medical University, Harbin, 150081 China

\*Correspondence: konglykl@sina.com

Received January 5, 2016 / Accepted February 29, 2016

The objective of this study was to investigate the key genes and pathways associated with thyroid carcinoma. Based on the microarray data of GSE27155, we identified the differentially expressed genes (DEGs) between four types of thyroid carcinoma samples (papillary carcinoma (PTC), oncocytic carcinoma (OTC), follicular carcinoma (FTC) and anaplastic carcinoma (ATC)) and normal controls. With the obtained DEGs, we performed gene functional interaction (FI) network analysis. Then we conducted Venn diagram analysis to identify the intersection and specific DEGs of the four types of thyroid carcinomas. The intersections DEGs were performed by functional enrichment and transcription factor (TF) prediction analyses. These specific DEGs were performed by pathway enrichment analysis. There were respectively 323, 318, 118 and 1005 DEGs identified in PTC, OTC, FTC and ATC. Twelve sub-network modules were extracted based on gene FI network analysis and eight thyroid carcinoma-associated DEGs were involved in the network, such as *TIMP1*. Based on the Venn diagram analysis, 27 common DEGs were identified, such as *HMGB3* which was regulated by TF of NKX3-1. There were 149 PTC-specific DEGs (like *CLDN1*), 160 OTC-specific DEGs, 94 FTC-specific DEGs (like *PPARG*), and 789 ATC-specific DEGs (like *CDK1*). They were enriched in some pathways, such as Cell cycle, Citrate cycle, and Oxidative phosphorylation. *TIMP1*, *HMGB3*, *CLDN1*, *CDK1* and *PPARG* as well as pathways of Cell cycle, Citrate cycle, and Oxidative phosphorylation may play important roles in the progression of thyroid carcinoma.

*Key words: thyroid carcinoma, differentially expressed gene, gene functional interaction network, transcription factor, pathway*

Thyroid carcinoma is the most common cancer of the endocrine system, accounting for approximately 1.7% of total cancer diagnoses [1]. It is estimated that more than 62,000 new cases of thyroid carcinoma were diagnosed in 2014 globally, which resulted in 1,890 deaths [2]. Thyroid carcinoma is classified according to its histopathological characteristics. Papillary thyroid carcinoma (PTC) is the

most prevalent thyroid carcinoma, which accounts for 80% of the cases [3]. Usually, PTC can be effectively treated with thyroidectomy and presents good prognosis [4]. Anaplastic thyroid carcinoma (ATC) is undifferentiated tumor of the thyroid follicular epithelium and accounts for less than 5% of all thyroid carcinomas [5]. ATC has a poor prognosis because of its extremely aggressive behavior [6]. Additionally, there are

**Abbreviations:** PTC – papillary thyroid carcinoma; ATC – anaplastic thyroid carcinoma; FTC – follicular thyroid carcinoma; MTC – medullary thyroid carcinoma; OTC – oncocytic thyroid carcinoma; GEO – Gene Expression Omnibus; DEG – differentially expressed gene; FI – functional interaction; TF – transcription factor; RMA – Robust Multi-array Averaging; FDR – false discovery rate; FC – fold change; CGI – cancer gene index; GO – Gene Ontology; DAVID – Database for Annotation, Visualization and Integrated Discovery; BP – biological process; CC – cellular component; KEGG – Kyoto Encyclopedia of Genes and Genomes; LSM7 – LSM7 homolog, U6 small nuclear RNA associated;

NNMT – nicotinamide N-methyltransferase; TIMP1 – TIMP metalloproteinase inhibitor 1; MYC – v-myc avian myelocytomatosis viral oncogene homolog; MYC – V-myc avian myelocytomatosis viral oncogene homolog; CHD3 – cadherin 3, type 1, P-cadherin (placental); DCN – decorin; DUSP1 – dual specificity phosphatase 1; FOS – FBJ murine osteosarcoma viral oncogene homolog; HMGB3 – mobility group-box 3; IL17B – interleukin 17B; NKX3-1 – NK3 homeobox 1; MEF2A – myocyte enhancer factor 2A; MMP – matrix metalloproteinase; PDHA1 – pyruvate dehydrogenase (lipoamide) alpha 1; ATP5B – ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide

another types of thyroid carcinomas, such as follicular thyroid carcinoma (FTC), medullary thyroid carcinoma (MTC), and oncocytic thyroid carcinoma (OTC) [5]. Iodine-deficient diets and environmental exposure to ionizing radiation are suspected to play a significant role in the etiology of thyroid carcinoma, but the underlying molecular mechanisms is still incompletely understood [7].

Recently, some genetic predisposing factors have been thought to be associated with the pathogenesis of thyroid carcinoma. For instance, mutation in RET proto-oncogene is associated with the development of thyroid carcinoma [8]. The dysregulations of prominin 1, and dickkopf WNT signaling pathway inhibitor 4 have been demonstrated in MTC [9]. Additionally, Garg et al. [6] suggested that laminin, gamma 2 played an important role in cell cycle, cell growth, migration and invasion of ATC cell line. Furthermore, overexpression of lysyl oxidase-like 2 is implicated with an increased number of apoptotic nuclei in MTC via involving in the cell adhesion pathway [9]. Activating mutations of several genes in the RET/RAS/BRAF/MAPK signal transduction pathway have been identified in PTC [10]. Although a number of genes and pathways have been suggested to influence the development of thyroid carcinoma, significant uncertainty remains regarding its molecular mechanisms.

Gene expression profiles are now providing great potential to clarify various aspects of carcinogenesis by revealing distinctive gene expression patterns related to various molecular and clinical parameters [11]. Therefore, in the present study, we downloaded the expression profile data of GSE27155 [12] from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database to identify the differentially expressed genes (DEGs) between four types of thyroid carcinoma samples and normal controls. With the obtained DEGs, we performed gene functional interaction (FI) network analysis to explore the FI network among DEGs. In order to study the common and specific DEGs of the 4 types of thyroid carcinoma, we conducted venn diagram analysis. Besides, we performed functional enrichment analysis and regulatory transcription factor (TF) prediction for these DEGs to investigate the key pathways and TFs. These DEGs and pathways would play roles in the the pathogenesis of thyroid carcinoma..

## Materials and methods

**Microarray data.** The microarray data of GSE27155 contributed by Giordano et al. [12] were downloaded from GEO database in National Center of Biotechnology Information, and the platform was GPL96 [HG-U133A] Affymetrix Human Genome U133A Array (Affymetrix Inc., Santa Clara, California, USA). There were five types of thyroid carcinoma samples in the dataset including 51 PTCs, eight OTCs, 13 FTCs, four ATCs, and two MTCs. Additionally, there were four normal thyroid samples. Among these thyroid carcinoma samples, MTCs were not used for analysis because of their

small sample size. Therefore, the other four types of thyroid carcinoma samples as well as normal thyroid samples were selected for further analysis.

**Data preprocessing.** The raw data in CEL format were downloaded and preprocessed using Robust Multi-array Averaging (RMA) method in affy (<http://www.bioconductor.org/packages/release/bioc/html/affy.html>) [13] package in Bioconductor (<https://www.bioconductor.org/>). The preprocessing processes included background correction, quantile normalization, and probe summarization. Combining with annotation packages of org.Hs.eg.db [14], hgu133a.db [15] and annotate [16] on the platform of GPL96 [HG-U133A] Affymetrix Human Genome U133A Array, the probe IDs were corresponded to gene symbols, obtaining 12434 gene expression matrixes.

**DEGs analysis.** The DEGs between four types of thyroid carcinoma samples and normal thyroid samples were analyzed using limma (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>) [17] package in R Bioconductor. Student's t-test was used to obtain the significant *p*-value of the DEGs. Then the *p*-value was adjusted as false discovery rate (FDR) with Benjamini & Hochberg [18] method. Additionally, the fold change (FC) was calculated.  $FDR < 0.05$  and  $|\log_2 FC| \geq 1$  were considered as the cut-off values.

Additionally, we investigated the four types of thyroid carcinoma-associated genes in Comparative Toxicogenomics Database (CTD, <http://ctdbase.org/>) [19]. CTD contains 24 million toxicogenomic connections between chemicals/drugs, genes/proteins and human diseases, which are manually curated from the scientific literature.

**Gene FI network analysis for DEGs.** ReactomeFIViz (available at [http://wiki.reactome.org/index.php/Reactome\\_FI\\_Cytoscape\\_Plugin](http://wiki.reactome.org/index.php/Reactome_FI_Cytoscape_Plugin)) [20] is a Cytoscape (<http://apps.cytoscape.org/apps/reactomefiviz>) [21] app for pathway and network-based data analysis based on the highly reliable Reactome FI network. In this study, ReactomeFIViz was used to analyze the gene FI network of the DEGs obtained from four types of thyroid carcinomas. In detail, the expression matrixes of all DEGs were input into the app. The correlations among DEGs involved in the same FIs were calculated, which were used as weights for edges in the whole FI network. Then Monte Carlo Localization graph clustering algorithm was applied to the weighted FI network to generate sub-network modules. Subsequently, the DEGs involved in the modules were performed pathway enrichment analysis ( $FDR < 0.05$ ). Finally, the cancer-related genes in the gene FI network were analyzed based on cancer gene index (CGI, <https://wiki.nci.nih.gov/display/cageneindex>) data of ReactomeFIViz. The CGI data contain annotations for cancer-related genes, and these annotations were extracted by using text-mining technologies and then validated by human curators (<https://wiki.nci.nih.gov/display/cageneindex/Creation+of+the+Cancer+Gene+Index>) [20].

**Venn diagram analysis for DEGs.** Based on the up- and down-regulated DEGs obtained in four types of thyroid car-

cinomas, we performed venn diagram analysis to identify the common (intersection) and specific DEGs of the 4 types of thyroid carcinomas using the online tool of Venny2.0 (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>).

**Functional enrichment analysis and TF prediction for intersection DEGs.** Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) [22] is an online tool aimed to provide functional explanations for a large number of genes derived from genomic studies and proteins. In this study, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses for the up- and down-regulated intersection DEGs in the venn diagram using DAVID. With the enrichment thresholds of  $p$ -value  $< 0.05$  and gene count  $\geq 2$ , the significant GO terms and KEGG pathways were obtained.

iRegulon (<http://iregulon.aertslab.org/>) [23] is a Cytoscape plugin integrating several TF databases, such as Transfac [24], Jaspar [25], Encode [26], Swissregulon [27] and Homer [28], which is used to detect enriched TF motifs and their optimal sets of direct targets by implementing a genome-wide ranking-and-recovery approach. At present, we predicted the TFs of intersection DEGs using the iRegulon Cytoscape plugin, and the TF-target gene interaction pairs with Normalized Enrichment Score (NES)  $> 4.0$  were selected.

**Pathway enrichment analysis for cancer-specific DEGs.** The DEGs specific to each type of thyroid carcinoma in the Venn diagram were performed KEGG pathway enrichment analysis using DAVID with thresholds of  $p$ -value  $< 0.05$  and gene count  $\geq$

## Results

**Identification of DEGs.** Comparing with normal controls, a total of 323 DEGs (220 up- and 103 down-regulated) were identified in PTC; 318 (201 up- and 117 down-regulated) in OTC; 118 (119 up- and 69 down-regulated) in FTC; 1005 (674 up- and 331 down-regulated) in ATC. Hierarchical clustering analysis of these DEGs and samples showed that these DEGs could differentiate cancer samples and normal controls (Figure 1).

Furthermore, there were respectively 41 PTC-associated, one OTC-associated, 5571 FTC-associated, and 750 ATC-associated genes in CTD. Additionally, compared with the present study, no common gene was found in PTC and OTC, while 85 (45.21%) common genes were found in FTC and 128 (12.74%) common genes were found in ATC.

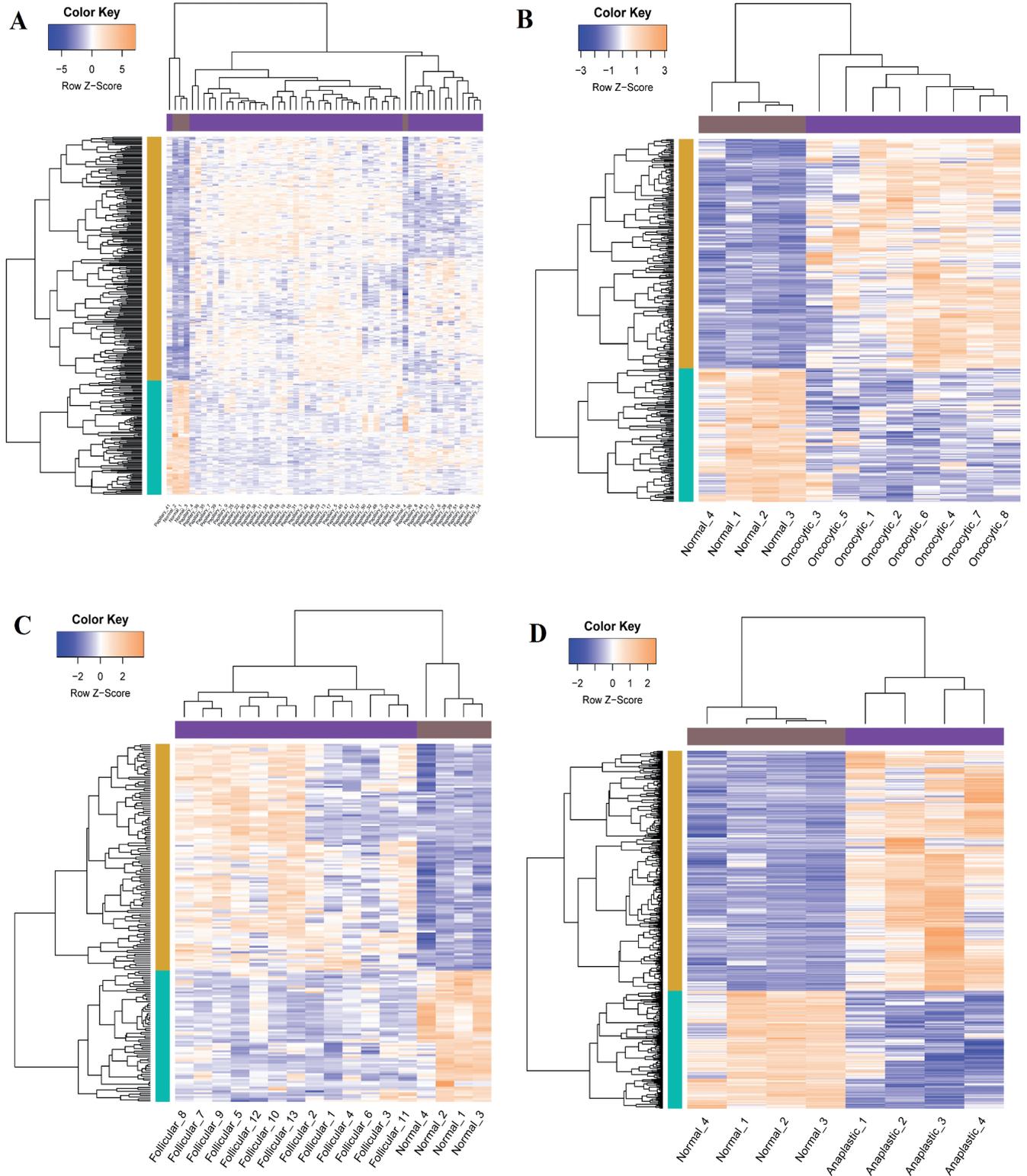
**Gene FI network analysis.** A total of 12 sub-network modules were obtained by ReactomeFIViz, which was shown in Figure 2. The network consisted of 193 genes and 59 FI pairs. The significant pathways enriched by DEGs in 12 modules were shown in Table 1. Additionally, 22 endocrine cancer-related DEGs were selected from CGI data, because thyroid carcinoma was a kind of endocrine-related cancer. Among these endocrine cancer-related DEGs, eight were associated

**Table 1. The significant pathways enriched by the differentially expressed genes (DEGs) involved in 12 sub-network modules.**

Module	GeneSet	Protein from module	FDR
0	Mitotic Metaphase and Anaphase (R)	14	5.00E-04
0	Mitotic Prometaphase (R)	17	5.00E-04
0	APC/C-mediated degradation of cell cycle proteins (R)	9	3.33E-04
1	HIF-1-alpha transcription factor network (N)	3	3.00E-02
1	Validated transcriptional targets of AP1 family members Fra1 and Fra2 (N)	2	1.22E-01
1	mechanism of gene regulation by peroxisome proliferators via ppara (B)	2	1.22E-01
2	GPCR downstream signaling (R)	14	1.00E-03
2	GPCR ligand binding (R)	11	5.00E-04
2	Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway (P)	8	3.33E-04
3	Calcineurin-regulated NFAT-dependent transcription in lymphocytes (N)	2	2.11E-01
3	Validated transcriptional targets of TAp63 isoforms (N)	2	1.14E-01
3	Transcriptional regulation of white adipocyte differentiation (R)	2	1.83E-01
4	AP-1 transcription factor network (N)	4	1.00E-03
4	IL6-mediated signaling events (N)	3	5.00E-04
4	ATF-2 transcription factor network (N)	3	6.67E-04
5	tumor suppressor arf inhibits ribosomal biogenesis (B)	2	1.80E-02
5	C-MYC pathway (N)	2	1.50E-02
5	Signaling by NOTCH1 (R)	2	6.50E-02
6	Cadherin signaling pathway (P)	5	1.00E-03
6	Wnt signaling pathway (P)	6	5.00E-04
6	Apoptotic execution phase (R)	3	1.33E-03
6	Signaling events mediated by VEGFR1 and VEGFR2 (N)	2	1.19E-01
7	Carbon metabolism (K)	6	1.00E-03
7	Biosynthesis of amino acids (K)	5	5.00E-04
7	2-Oxocarboxylic acid metabolism (K)	3	3.33E-04
8	Metabolism of carbohydrates (R)	6	1.00E-03
8	Glycosaminoglycan biosynthesis – chondroitin sulfate / dermatan sulfate (K)	3	5.00E-04
8	Endogenous TLR signaling (N)	2	1.13E-02
8	Extracellular matrix organization (R)	3	2.35E-02
9	Signaling by Rho GTPases (R)	6	1.00E-03
9	IL8- and CXCR2-mediated signaling events (N)	2	3.25E-02
10	Beta1 integrin cell surface interactions (N)	7	1.00E-03
10	Protein digestion and absorption (K)	6	5.00E-04
10	Integrin signalling pathway (P)	6	3.33E-04
11	Processing of Capped Intron-Containing Pre-mRNA (R)	7	1.00E-03
11	Metabolism of non-coding RNA (R)	5	5.00E-04
11	Processing of Capped Intronless Pre-mRNA (R)	3	3.33E-04

FDR: false discovery rate.

Pathway database: R – Reactome, K – KEGG, N – NCI PID, P – Panther, and B – BioCarta.



**Figure 1.** Heatmap plot of differentially expressed genes (DEGs) across all samples. Upper color bar represents sample classes: purple represents thyroid carcinoma group and brown represents control group. A: papillary thyroid carcinoma; B: oncocytic thyroid carcinoma; C: follicular thyroid carcinoma; D: anaplastic thyroid carcinoma.

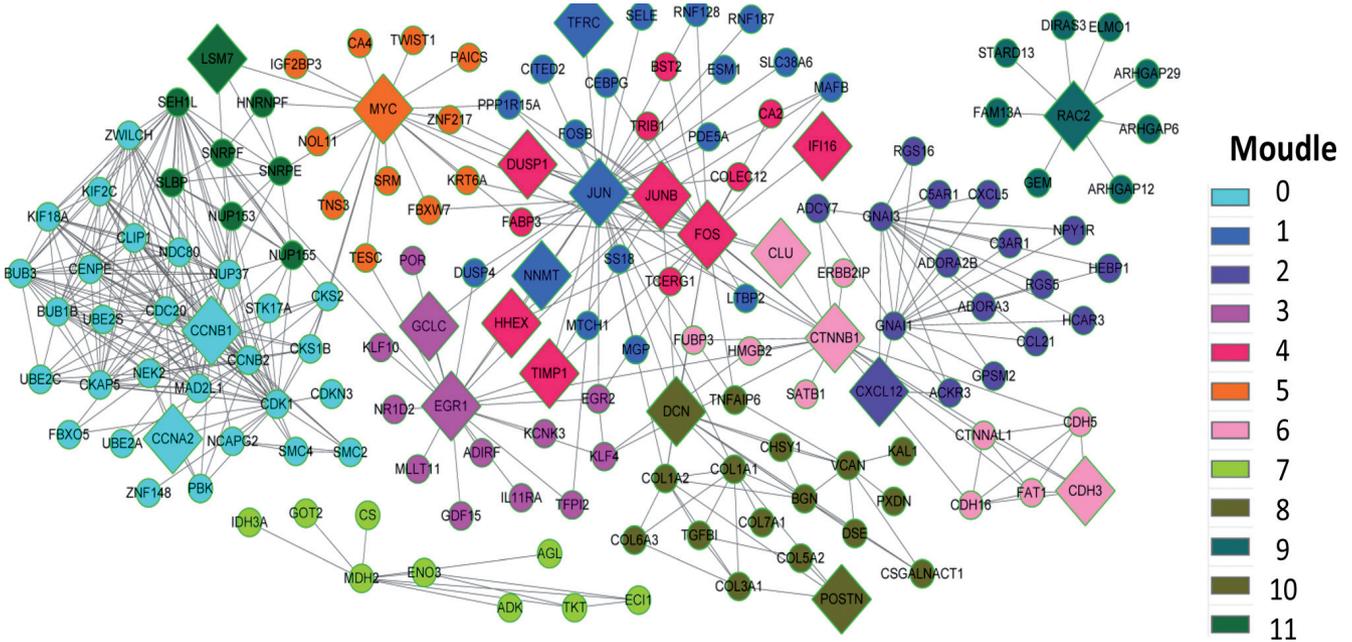
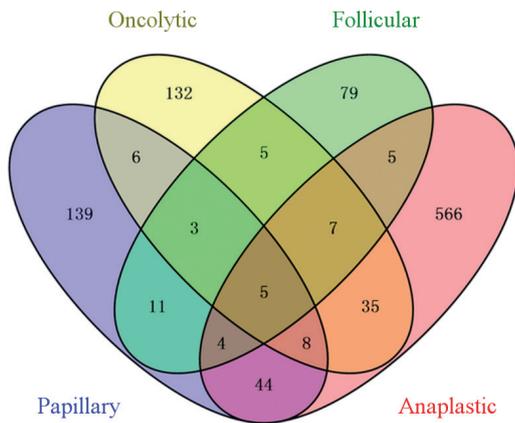


Figure 2. Sub-network modules obtained by gene functional interaction network analysis. Each color represents one module. Diamond represents gene node that is associated with endocrine carcinoma.

with thyroid carcinoma, including LSM7 homolog, U6 small nuclear RNA associated (*LSM7*), nicotinamide N-methyltransferase (*NNMT*), TIMP metalloproteinase inhibitor 1 (*TIMP1*), v-myc avian myelocytomatosis viral oncogene homolog (*MYC*), cadherin 3, type 1, P-cadherin (placental) (*CDH3*), decorin (*DCN*), dual specificity phosphatase 1 (*DUSP1*) and FBJ murine osteosarcoma viral oncogene homolog (*FOS*).

**Venn diagram analysis.** The Venn diagram of DEGs was shown in Figure 3. The results showed that there were five common (intersection) up-regulated DEGs and 22 common down-regulated DEGs for the four types of thyroid carcinomas, such as high mobility group-3 (*HMGB3*), and interleukin 17B (*IL17B*). In addition, there were 139 up- and ten down-regulated PTC-specific DEGs (like claudin-1

**A** Up-regulated genes in four types for thyroid carcinoma



**B** Down-regulated genes in four types for thyroid carcinoma

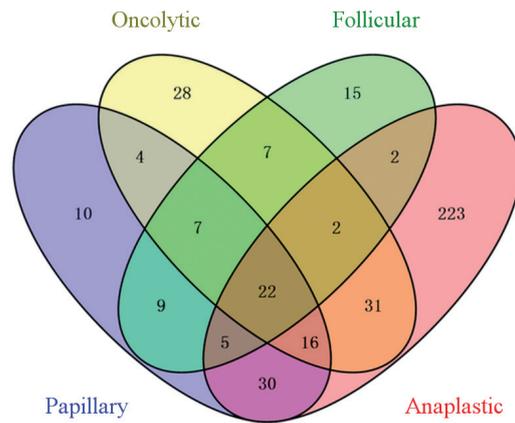


Figure 3. The Venn diagram for up-regulated differentially expressed genes (DEGs) (A) and down-regulated DEGs (B). Purple represents papillary thyroid carcinoma; Yellow represents oncocytic thyroid carcinoma; Green represents follicular thyroid carcinoma; Pink represents anaplastic thyroid carcinoma.

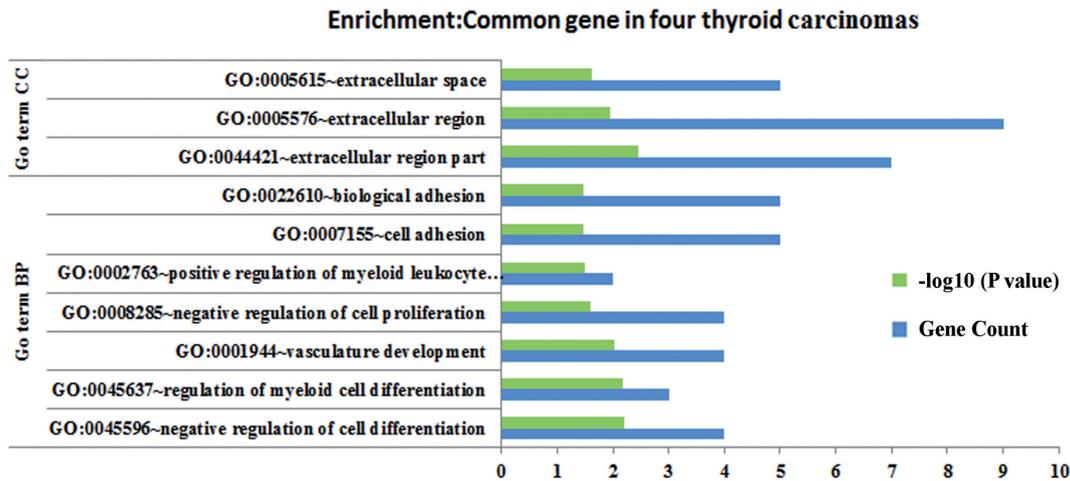


Figure 4. The cellular component (CC) and biological process (BP) functions enriched by the common genes in four thyroid carcinomas.

(*CLDN1*), 132 up- and 28 down-regulated OTC-specific DEGs, 79 up- and 15 down-regulated FTC-specific DEGs (such as proliferator-activated receptor gamma (*PPARG*)), and 566 up- and 223 down-regulated ATC-specific DEGs (such as cyclin-dependent kinase 1 (*CDK1*)).

**Functional enrichment analysis and TF prediction for intersection DEGs.** The intersection DEGs for the four types of thyroid carcinomas were mainly enriched in ten GO terms, including three cellular component (CC) related terms (such as extracellular space) and seven biological process (BP) related terms (such as negative regulation of cell differentiation) (Figure 4). However, no pathway was enriched by these DEGs. Furthermore, nine TFs, such as NK3 homeobox 1 (*NKX3-1*),

and myocyte enhancer factor 2A (*MEF2A*), were predicted for the 27 intersection DEGs, and the transcriptional regulatory network was shown in Figure 5.

**Functional enrichment analysis for cancer specific DEGs.** The pathways enriched by the DEGs specific to each type of thyroid carcinoma were shown in Table 2. However, among the down-regulated specific DEGs, only ATC-specific down-regulated DEGs were enriched in pathways. PTC-specific up-regulated DEGs were significantly enriched in pathways of hsa04514:Cell adhesion molecules (CAMs), and hsa04610:Complement and coagulation cascades. OTC-specific up-regulated DEGs were mainly enriched in hsa00020:Citrate cycle (TCA cycle), and hsa00190:Oxidative

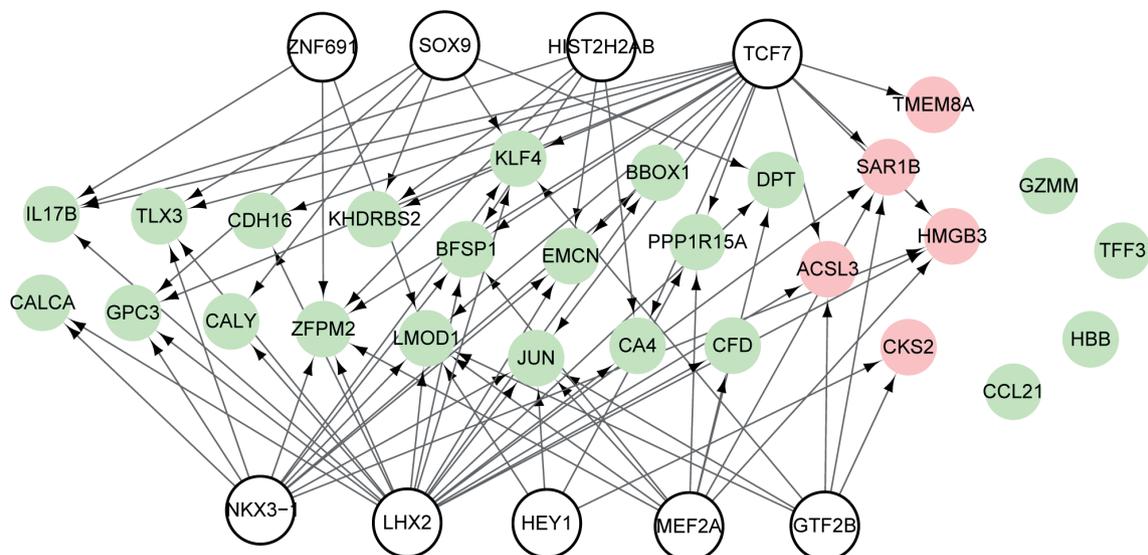


Figure 5. The transcriptional regulatory network of common genes and their predicted transcription factors. White represents transcription factor; Pink represents up-regulated differentially expressed genes (DEGs); Green represents down-regulated DEGs.

**Table 2. The pathways (top 5) enriched by the differentially expressed genes (DEGs) specific to each type of thyroid carcinoma.**

	Category	Term	Count	P-value
		<b>Papillary</b>		
up	KEGG_PATHWAY	hsa04514:Cell adhesion molecules (CAMs)	7	0.003654
	KEGG_PATHWAY	hsa04610:Complement and coagulation cascades	5	0.007596
	KEGG_PATHWAY	hsa04530:Tight junction	6	0.017679
	KEGG_PATHWAY	hsa04670:Leukocyte transendothelial migration	5	0.044768
		<b>Oncocytic</b>		
up	KEGG_PATHWAY	hsa05012:Parkinson's disease	18	3.41E-13
	KEGG_PATHWAY	hsa05010:Alzheimer's disease	19	1.60E-12
	KEGG_PATHWAY	hsa00020:Citrate cycle (TCA cycle)	11	3.28E-12
	KEGG_PATHWAY	hsa05016:Huntington's disease	19	8.88E-12
	KEGG_PATHWAY	hsa00190:Oxidative phosphorylation	16	8.13E-11
		<b>Follicular</b>		
up	KEGG_PATHWAY	hsa03320:PPAR signaling pathway	4	0.006145
	KEGG_PATHWAY	hsa04910:Insulin signaling pathway	4	0.036947
		<b>Anaplastic</b>		
up	KEGG_PATHWAY	hsa05130:Pathogenic Escherichia coli infection	15	2.33E-07
	KEGG_PATHWAY	hsa04110:Cell cycle	19	1.84E-05
	KEGG_PATHWAY	hsa04512:ECM-receptor interaction	13	5.07E-04
	KEGG_PATHWAY	hsa03030:DNA replication	8	0.001231
	KEGG_PATHWAY	hsa04540:Gap junction	12	0.002871
down	KEGG_PATHWAY	hsa00650:Butanoate metabolism	7	8.81E-06
	KEGG_PATHWAY	hsa00330:Arginine and proline metabolism	7	1.20E-04
	KEGG_PATHWAY	hsa00380:Tryptophan metabolism	5	0.002802
	KEGG_PATHWAY	hsa00280:Valine, leucine and isoleucine degradation	5	0.003979
	KEGG_PATHWAY	hsa00410:beta-Alanine metabolism	4	0.004028

phosphorylation. Two pathways were enriched by the up-regulated DEGs specific to FTC, including hsa03320:PPAR signaling pathway, and hsa04910:Insulin signaling pathway. The up-regulated specific DEGs of ATC were mainly enriched in hsa04110:Cell cycle, and hsa04512:ECM-receptor interaction, while the down-regulated DEGs were mainly enriched in metabolism-associated pathways.

## Discussion

The thyroid is an inconspicuous organ, however the morbidity of thyroid is much higher than the other endocrine organs [29]. Importantly, the underlying molecular mechanisms of thyroid carcinoma are still largely unknown. In the present study, hundreds of DEGs were identified between each type of thyroid carcinoma samples and normal controls, respectively. Twelve sub-network modules were extracted based on gene FI network analysis for these DEGs and eight thyroid carcinoma-associated DEGs were involved in the network. Based on the venn diagram analysis, the common DEGs and specific DEGs of the four types of thyroid carcinoma were identified, besides, they were enriched in many GO terms and pathways.

During cancer progression, cancer cells can modify their surrounding host microenvironment largely through activities

of the MMPs which play important roles in tumor invasion and metastasis [30]. Importantly, the bioactivities of MMPs are regulated by the tissue inhibitors of metalloproteinases (TIMPs) [31]. Interestingly, our results showed that *TIMP1* was a DEG associated with thyroid carcinoma in the gene FI network. *TIMP1* belongs to the TIMP gene family, the proteins encoded by which are natural inhibitors of the matrix metalloproteinases (MMPs), and are able to promote cell proliferation in a wide range of cell types [32]. Increased expression of *TIMP1* has been found to be associated with worse prognosis of various tumors, such as laryngeal carcinoma and melanoma [33]. Importantly, *TIMP1* up-regulation in PTC was reported recently [34]. Earlier studies indicated that TIMP1 displayed anti-cancer activities [35, 36], however, recent studies demonstrated a paradoxical protumor effect of TIMP1 [31, 37]. In the present study, *TIMP1* was up-regulated in PTC and ATC, while down-regulated in FTC, which was in accordance with the findings above. Taken together, we speculated that *TIMP1* might be used as a biomarker for PTC, ATC and FTC.

The venn diagram analysis found 27 common DEGs for the four types of thyroid carcinomas, such as *HMGB3*. *HMGB3*, encoding a member of high mobility group (HMG) box subfamily, plays an important role in maintaining stem cell populations, and may be aberrantly expressed in tumor cells. Two members of HMG-Box subfamily, *HMGB1* and *HMGB2*,

have been suggested to play important roles in cell proliferation, invasion and metastasis of cancer progression [38]. Recent study of Nemeth et al. [39] showed that HMGB3 played a key role in regulating the balance between proliferation and differentiation in the primitive stages of hematopoiesis. Overexpression of *HMGB3* has been found in metastatic breast cancers [40]. In this study, *HMGB3* was significantly enriched in GO terms associated with negative regulation of cell differentiation. Additionally, it was regulated by several TFs, such as NKX3-1. This TF functions as a negative regulator of epithelial cell growth in prostate tissue, aberrant expression of which is associated with prostate tumor progression [41]. There is hardly any study about the role of NKX3-1 in thyroid carcinoma, thus, we inferred that NKX3-1 might play a role in thyroid carcinoma by regulating *HMGB3*. Additionally, *HMGB3* might play an important role in the progression of the four types of thyroid carcinomas through regulation of cell differentiation.

In addition to common DEGs, there were a lot of DEGs specific to each type of thyroid carcinoma, besides, they were enriched in several pathways. Among the four types of thyroid carcinoma, PTC is the most prevalent, and ATC is the most malignant [3]. *CLDN1* was a PTC-specific DEG, which was enriched in the pathway of hsa04530:Tight junction. Bhat et al. [42] have reported that *CLDN1* is a key constituent of tight junctions which are implicated with many human diseases, including cancer. Additionally, *CLDN1* was up-regulated in PTC, which was in keep with previous studies: *CLDN1* was significantly overexpressed in many cancer and associated with tumor growth and progression [43, 44]. Therefore, the up-regulation of *CLDN1* might play a key role in PTC progression. *CDK1*, in this study, was a ATC-specific DEG, and it was enriched in hsa04110:Cell cycle. As is well known, deregulation of cell cycle is a common feature of cancer. In mammalian, cell cycle is controlled by a subfamily of cyclin-dependent kinases (CDKs), dysregulation of which induce unscheduled proliferation [45]. *CDK1* is a member of the CDK family, besides, it is the only essential cell cycle CDK in human cells [46]. As a result, *CDK1* might involve in the development of ATC through the pathway of cell cycle.

Mitochondria are essential organelles and have been suggested to play vital roles in controlling cancer cells fate and proliferation. Importantly, mitochondria regulate tumor growth through modulating the citrate cycle (TCA cycle) and oxidative phosphorylation [47]. In the present study, several OTC-specific DEGs were significantly enriched in the pathways of hsa00020:Citrate cycle (TCA cycle) and hsa00190:Oxidative phosphorylation, such as pyruvate dehydrogenase (lipoamide) alpha 1 (*PDHA1*) and ATP synthase, H<sup>+</sup> transporting, mitochondrial F1 complex, beta polypeptide (*ATP5B*). The 2 DEGs have been found to be associated with cancers [48, 49]. Furthermore, FTC-specific DEGs were only significantly involved in 2 pathways. hsa03320:PPAR signaling pathway, for instance, was enriched by *PPARG*. *PPARG*

encodes a member of the PPAR subfamily of nuclear receptors that involve in many important physiological processes, including lipid metabolism, cell differentiation, and apoptosis [50, 51]. Studies have revealed that *PPARG* is generally up-regulated in liposarcoma, colon, and breast cancer [52, 53]. Presently, its role in FTC has not been reported, but we think that it may be used as a biomarker of PTC.

In summary, our study used bioinformatic methods to identify the genes and pathways associated with thyroid carcinoma. These DEGs, such as *TIMP1*, *HMGB3*, *CLDN1*, *CDK1*, and *PPARG* as well as the TF of NKX3-1 may have the potential to be used as targets for the diagnosis and treatment of thyroid carcinoma. Additionally, these pathways, including Tight junction, Cell cycle, Citrate cycle (TCA cycle), and Oxidative phosphorylation may also play important roles in the progression of thyroid carcinoma. Moreover, findings of the present study may contribute to the further understanding of the underlying molecular mechanisms of thyroid carcinoma. However, further experimental studies are still needed to confirm the results of this study.

## References

- [1] HE W, QI B, ZHOU Q, LU C, HUANG Q, et al. Key genes and pathways in thyroid cancer based on gene set enrichment analysis. *Oncol Rep* 2013; 30: 1391–1397.
- [2] SIEGEL R, MA J, ZOU Z, JEMAL A. Cancer statistics, 2014. *CA Cancer J Clin* 2014; 64: 9–29. <http://dx.doi.org/10.3322/caac.21208>
- [3] GERALDO MV, KIMURA ET. Integrated Analysis of Thyroid Cancer Public Datasets Reveals Role of Post-Transcriptional Regulation on Tumor Progression by Targeting of Immune System Mediators. *PLoS One* 2015; 10: e0141726. <http://dx.doi.org/10.1371/journal.pone.0141726>
- [4] DEGROOT LJ, KAPLAN EL, MCCORMICK M, STRAUS FH. Natural History, Treatment, and Course of Papillary Thyroid Carcinoma\*. *J Clin Endocrinol Metab* 1990; 71: 414–424. <http://dx.doi.org/10.1210/jcem-71-2-414>
- [5] MITCHELL RS, KUMAR V, ABBAS AK, FAUSTO N (2007). *Robbins Basic Pathology*. Philadelphia: Saunders. Editor (ed)^(eds). ISBN 1–4160–2973–7: City.
- [6] GARG M, KANOJIA D, OKAMOTO R, JAIN S, MADAN V, et al. Laminin-5 $\gamma$ -2 (LAMC2) is highly expressed in anaplastic thyroid carcinoma and is associated with tumor progression, migration, and invasion by modulating signaling of EGFR. *J Clin Endocrinol Metab* 2013; 99: E62–E72. <http://dx.doi.org/10.1210/jc.2013-2994>
- [7] RODRIGUEZ-RODERO S, FERNANDEZ AF, FERNÁNDEZ-MORERA JL, CASTRO-SANTOS P, BAYON GF, et al. DNA methylation signatures identify biologically distinct thyroid cancer subtypes. *J Clin Endocrinol Metab* 2013; 98: 2811–2821. <http://dx.doi.org/10.1210/jc.2012-3566>
- [8] TACCALITI A, SILVETTI F, PALMONELLA G, BOSCARO M. Genetic alterations in medullary thyroid cancer: diagnostic and prognostic markers. *Curr Genomics* 2011; 12: 618. <http://dx.doi.org/10.2174/138920211798120835>

- [9] MALISZEWSKA A, LEANDRO-GARCIA LJ, CASTEL-BLANCO E, MACIÀ A, DE CUBAS A, et al. Differential gene expression of medullary thyroid carcinoma reveals specific markers associated with genetic conditions. *Am J Pathol* 2013; 182: 350–362. <http://dx.doi.org/10.1016/j.ajpath.2012.10.025>
- [10] CAPELLA G, MATIAS-GUIU X, AMPUDIA X, DE LEIVA A, PERUCHO M, et al. Ras oncogene mutations in thyroid tumors: polymerase chain reaction-restriction-fragment-length polymorphism analysis from paraffin-embedded tissues. *Diagn Mol Pathol* 1996; 5: 45–52. <http://dx.doi.org/10.1097/00019606-199603000-00008>
- [11] HANASH S. Integrated global profiling of cancer. *Nat Rev Cancer* 2004; 4: 638–644. <http://dx.doi.org/10.1038/nrc1414>
- [12] GIORDANO TJ, KUICK R, THOMAS DG, MISEK DE, VINCO M, et al. Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* 2005; 24: 6646–6656. <http://dx.doi.org/10.1038/sj.onc.1208822>
- [13] GAUTIER L, COPE L, BOLSTAD BM, IRIZARRY RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; 20: 307–315. <http://dx.doi.org/10.1093/bioinformatics/btg405>
- [14] CARLSON M. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.2.3. 2015.
- [15] CARLSON M. hgu133a.db: Affymetrix Human Genome U133 Set annotation data (chip hgu133a). R package version 3.2.2.
- [16] GENTLEMAN R. annotate: Annotation for microarrays. R package version 1.48.0.
- [17] RITCHIE ME, PHIPSON B, WU D, HU Y, LAW CW, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; gkv007. <http://dx.doi.org/10.1093/nar/gkv007>
- [18] BENJAMINI Y, DRAID, ELMER G, KAFKAFIN, GOLANI I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001; 125: 279–284. [http://dx.doi.org/10.1016/S0166-4328\(01\)00297-2](http://dx.doi.org/10.1016/S0166-4328(01)00297-2)
- [19] DAVIS AP, GRONDIN CJ, LENNON-HOPKINS K, SARACENI-RICHARDS C, SCIAKY D, et al. The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. *Nucleic acids research* 2015; 43: D914–D920. <http://dx.doi.org/10.1093/nar/gku935>
- [20] WU G, DAWSON E, DUONG A, HAW R, STEIN L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. *F1000Res* 2014; 3: 146.
- [21] SHANNON P, MARKIEL A, OZIER O, BALIGA NS, WANG JT, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498–2504. <http://dx.doi.org/10.1101/gr.1239303>
- [22] ALVORD G, ROAYAEI J, STEPHENS R, BASELER MW, LANE HC, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 2007; 8: 183. <http://dx.doi.org/10.1186/gb-2007-8-9-r183>
- [23] REKIN'S JANKY AV, HANA IMRICOVA, BRAM VAN DE SANDE, LAURA STANDAERT, VALERIE CHRISTIAENS, GERT HULSELMANS, KOEN HERTEN, MARINA NAVAL SANCHEZ, DELPHINE POTIER, DMITRY SVETLICHNYY ZKA, MARK FIERS, JEAN-CHRISTOPHE MARINE, STEIN AERTS. iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. *PLoS Comput Biol* 2014.
- [24] WINGENDER E, CHEN X, FRICKE E, GEFFERS R, HEHL R, et al. The TRANSFAC system on gene expression regulation. *Nucleic Acids Res* 2001; 29: 281–283. <http://dx.doi.org/10.1093/nar/29.1.281>
- [25] PORTALES-CASAMAR E, THONGJUEA S, KWON AT, ARENILLAS D, ZHAO X, et al. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2009; gkp950.
- [26] NEPH S, VIERSTRA J, STERGACHIS AB, REYNOLDS AP, HAUGEN E, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 2012; 489: 83–90. <http://dx.doi.org/10.1038/nature11212>
- [27] PACHKOV M, BALWIERZ PJ, ARNOLD P, OZONOV E, VAN NIMWEGEN E. SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates. *Nucleic Acids Res* 2013; 41: D214–D220. <http://dx.doi.org/10.1093/nar/gks1145>
- [28] HEINZ S, BENNER C, SPANN N, BERTOLINO E, LIN YC, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; 38: 576–589. <http://dx.doi.org/10.1016/j.molcel.2010.05.004>
- [29] WILSON C. Thyroid function: Thyroid dysfunction and risk of atrial fibrillation. *Nat Rev Endocrinol* 2013; 9: 63–63. <http://dx.doi.org/10.1038/nrendo.2012.239>
- [30] STAMENKOVIC I (2000). Matrix metalloproteinases in tumor invasion and metastasis, Vol. 10. Editor (ed)^(eds). Elsevier: City, pp. 415–433. <http://dx.doi.org/10.1006/scbi.2000.0379>
- [31] BOURBOULIA D, STETLER-STEVENSON WG (2010). Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion, Vol. 20. Editor (ed)^(eds). Elsevier: City, pp. 161–168. <http://dx.doi.org/10.1016/j.semcancer.2010.05.002>
- [32] HORNEBECK W. Down-regulation of tissue inhibitor of matrix metalloprotease-1 (TIMP-1) in aged human skin contributes to matrix degradation and impaired cell growth and survival. *Pathol Biol* 2003; 51: 569–573. <http://dx.doi.org/10.1016/j.patbio.2003.09.003>
- [33] TARHINI AA, LIN Y, YEKU O, LAFRAMBOISE WA, ASHRAF M, et al. A four-marker signature of TNF-RII, TGF- $\alpha$ , TIMP-1 and CRP is prognostic of worse survival in high-risk surgically resected melanoma. *J Transl Med* 2014; 12: 19. <http://dx.doi.org/10.1186/1479-5876-12-19>
- [34] ALESSANDRA B, PIERINA R, ELVIRA C, GIUSEPPE P, VITO R, et al. BRAFV600E mutation, TIMP-1 upregulation, and NF- $\kappa$ B activation: closing the loop on the papillary thyroid cancer trilogy. *Endocrine Related Cancer* 2011; 18: 669–685. <http://dx.doi.org/10.1530/ERC-11-0076>

- [35] GU P, XING X, TÄNZER M, RÖCKEN C, WEICHERT W, et al. Frequent loss of TIMP-3 expression in progression of esophageal and gastric adenocarcinomas. *Neoplasia* 2008; 10: 563–572. <http://dx.doi.org/10.1593/neo.08208>
- [36] DENG X, HE G, LEVINE A, CAO Y, MULLINS C. Adenovirus - mediated expression of TIMP - 1 and TIMP - 2 in bone inhibits osteolytic degradation by human prostate cancer. *Int J Cancer* 2008; 122: 209–218. <http://dx.doi.org/10.1002/ijc.23053>
- [37] SCHELTER F, GRANDL M, SEUBERT B, SCHATEN S, HAUSER S, et al. Tumor cell-derived Timp-1 is necessary for maintaining metastasis-promoting Met-signaling via inhibition of Adam-10. *Clin Exp Metastasis* 2011; 28: 793–802. <http://dx.doi.org/10.1007/s10585-011-9410-z>
- [38] KOSTOVA N, ZLATEVA S, UGRINOVA I, PASHEVA E. The expression of HMGB1 protein and its receptor RAGE in human malignant tumors. *Mol Cell Biochem* 2010; 337: 251–258. <http://dx.doi.org/10.1007/s11010-009-0305-0>
- [39] NEMETH MJ, KIRBY MR, BODINE DM. Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. *Proc Natl Acad Sci* 2006; 103: 13783–13788. <http://dx.doi.org/10.1073/pnas.0604006103>
- [40] ELGAMAL OA, PARK J-K, GUSEV Y, AZEVEDO-POULY A, JIANG J, et al. Tumor suppressive function of mir-205 in breast cancer is linked to HMGB3 regulation. *PLoS One* 2013; 8: e76402. <http://dx.doi.org/10.1371/journal.pone.0076402>
- [41] TAN PY, CHANG CW, CHNG KR, WANSA KD, SUNG WK, et al. Integration of regulatory networks by NKX3–1 promotes androgen-dependent prostate cancer survival. *Mol Cell Biol* 2012; 32: 399–414. <http://dx.doi.org/10.1128/MCB.05958-11>
- [42] BHAT AA, SHARMA A, POPE J, KRISHNAN M, WASHINGTON MK, et al. Caudal homeobox protein Cdx-2 cooperates with Wnt pathway to regulate claudin-1 expression in colon cancer cells. *PLoS One* 2012; 7: e37174. <http://dx.doi.org/10.1371/journal.pone.0037174>
- [43] SUH Y, YOON C, KIM R, LIM E, OH Y, et al. Claudin-1 induces epithelial–mesenchymal transition through activation of the c-Abl-ERK signaling pathway in human liver cells. *Oncogene* 2013; 32: 4873–4882. <http://dx.doi.org/10.1038/onc.2012.505>
- [44] GROSSE B, CASSIO D, YOUSEF N, BERNARDO C, JACQUEMIN E, et al. Claudin - 1 involved in neonatal ichthyosis sclerosing cholangitis syndrome regulates hepatic paracellular permeability. *Hepatology* 2012; 55: 1249–1259. <http://dx.doi.org/10.1002/hep.24761>
- [45] MARCOS M, MARIANO B. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* 2009; 9: 153–166. <http://dx.doi.org/10.1038/nrc2602>
- [46] BROWN NR, KOROLCHUK S, MARTIN MP, STANLEY WA, MOUKHAMETZIANOV R, et al. CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat Commun* 2015; 6. <http://dx.doi.org/10.1038/ncomms7769>
- [47] SOLAINI G, SGARBI G, BARACCA A. Oxidative phosphorylation in cancer cells. *Biochim Biophys Acta* 2011; 1807: 534–542. <http://dx.doi.org/10.1016/j.bbabbio.2010.09.003>
- [48] LIU F, ZHANG W, YOU X, LIU Y, LI Y, et al. The oncoprotein HBXIP promotes glucose metabolism reprogramming via downregulating SCO2 and PDHA1 in breast cancer. *Oncotarget* 2015; 6: 27199–27213. <http://dx.doi.org/10.18632/oncotarget.4508>
- [49] HJERPE E, BRAGE SE, CARLSON J, STOLT MF, SCHEDVINS K, et al. Metabolic markers GAPDH, PKM2, ATP5B and BEC-index in advanced serous ovarian cancer. *BMC Clin Pathol* 2013; 13: 30. <http://dx.doi.org/10.1186/1472-6890-13-30>
- [50] WILLSON T, BROWN P, STERNBACH D, BR. The PPARs: from orphan receptors to drug discovery. *J Med Chem* 2000; 43: 527–550. <http://dx.doi.org/10.1021/jm990554g>
- [51] ELROD HA, SUN SY. PPARgamma and Apoptosis in Cancer. *Ppar Research* 2008; 2008: -.
- [52] DUBOIS RN, GUPTA R, BROCKMAN J, REDDY BS, KRAKOW SL, et al. The nuclear eicosanoid receptor, PPARgamma, is aberrantly expressed in colonic cancers. *Carcinogenesis* 1998; 19: 49–53. <http://dx.doi.org/10.1093/carcin/19.1.49>
- [53] TONTONAZ P, SINGER S, FORMAN BM, SARRAF P, FLETCHER JA, et al. Terminal differentiation of human liposarcoma cells induced by ligands for peroxisome proliferator-activated receptor gamma and the retinoid X receptor. *Proceedings of the National Academy of Sciences of the United States of America* 1997; 94: 237–241. <http://dx.doi.org/10.1073/pnas.94.1.237>