Identification of Key Pathways and Genes in Anaplastic Thyroid Carcinoma via Integrated Bioinformatics Analysis

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEF **Shengqing Hu**
AG **Yunfei Liao**
AG **Lulu Chen**

Department of Endocrinology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, P.R. China

Corresponding Author: Lulu Chen, e-mail: cheria_chen@126.com
Source of support: This work was supported by grants 2016YFC0901200 and 2016YFC0901203 from the Ministry of Science and Technology and a grant from the National Natural Science Foundation of China (81471069)

Background: To provide a better understanding of anaplastic thyroid carcinoma (ATC) at the molecular level, this study aimed to identify the genes and key pathways associated with ATC by using integrated bioinformatics analysis.

Material/Methods: Based on the microarray data GSE9115, GSE65144, and GSE53072 derived from the Gene Expression Omnibus, the differentially expressed genes (DEGs) between ATC samples and normal controls were identified. With DEGs, we performed a series of functional enrichment analyses. Then, a protein–protein interaction (PPI) network was constructed and visualized, with which the hub gene nodes were screened out. Finally, modules analysis for the PPI network was performed to further investigate the potential relationships between DEGs and ATC.

Results: A total of 537 common DEGs were screened out from all 3 datasets, among which 247 genes were upregulated and 275 genes were downregulated. GO analysis indicated that upregulated DEGs were mainly involved in cell division and mitotic nuclear division and the downregulated DEGs were significantly enriched in ventricular cardiac muscle cell action potential. KEGG pathway analysis showed that the upregulated DEGs were mainly enriched in cell cycle and ECM-receptor interaction and the downregulated DEGs were mainly enriched in thyroid hormone synthesis, insulin resistance, and pathways in cancer. The top 10 hub genes in the constructed PPI network were CDK1, CCNB1, TOP2A, AURKB, CCNA2, BUB1, AURKA, CDC20, MAD2L1, and BUB1B. The modules analysis showed that genes in the top 2 significant modules of PPI network were mainly associated with mitotic cell cycle and positive regulation of mitosis, respectively.

Conclusions: We identified a series of key genes along with the pathways that were most closely related with ATC initiation and progression. Our results provide a more detailed molecular mechanism for the development of ATC, shedding light on the potential biomarkers and therapeutic targets.

MeSH Keywords: **Carcinoma • Gene Expression Profiling • Molecular Biology**

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/910088

📄 2532   ▦ 4   📊 5   📖 45

Background

Thyroid cancer is one of the most common types of endocrine malignancy, the incidence of which is increasing to be among the top malignancies in the United States [1]. It is predicted that approximately 63 000 new cases of thyroid cancer were diagnosed in 2014 compared with 37 200 in 2009 [2]. Anaplastic thyroid carcinoma (ATC), which accounts for about 1–2% of all thyroid carcinomas, is considered to be highly resistant to all therapies, with a 1-year survival rate of 5–30% [3] and has a poor prognosis due to its extremely aggressive behavior [4]. Accumulating evidence demonstrates the participation of multiple genes and cellular pathways in the occurrence and development of ATC. Understanding the molecular mechanism involved in proliferation, apoptosis, and invasion will greatly improve diagnosis and therapy. To date, however, little is known about the precise molecular mechanism underlying ATC progression.

In recent years, microarray technology, a high-throughput approach, has been widely used to profile gene expression [5]. Many studies have made use of microarray technology to explore the differentially expressed genes (DEGs) in distinct pathways, biological processes, molecular functions, or cellular components [6–8]. Although current microarray technology has limited detection sensitivity and dynamic range, which restricts its ability to measure changes in low-level gene expression [9], combining it with bioinformatics analysis enables comprehensive analysis of changes in expression of genes, providing a new and effective way to explore the molecular mechanisms of various cancers [10–12]. However, a reliable and accurate result with respect to ATC is still lacking.

Therefore, in this study, we aimed to identify the key pathways and genes in ATC with the aid of extensive bioinformatics methods. We obtained 3 mRNA expression profile data sets - GSE9115, GSE65144, and GSE53072 – to conduct collaborative microarray analysis. The DEGs were first screened between ATC tissues and normal tissues for each data set. Then, the common DEGs from these 3 data sets were selected for gene ontology (GO) and pathway enrichment analysis, followed by protein–protein interaction (PPI) network construction and module analysis. Our results provide further insights into the development of ATC at the molecular level and reveal potential diagnostic biomarkers and therapeutical targets for ATC.

Material and Methods

Microarray data

We obtained 3 gene expression profiles (GSE9115, GSE65144, and GSE53072) from the Gene Expression Omnibus (GEO, *http://www.ncbi.nlm.nih.gov/geo*), which is a public repository of gene expression datasets, original series, and platform records. The array data of GSE9115 includes 4 normal thyroid samples and 5 ATC samples [13]. GSE65144 consists of 13 normal thyroid samples and 12 ATC samples [14]. GSE53072 contains 4 normal thyroid samples and 5 ATC samples [15].

Identification of DEGs

We used statistical analysis software R (version 3.5.0, *https://www.r-project.org/*) and packages of Bioconductor (*http://www.bioconductor.org/*) to conduct significance analysis of DEGs between ATC samples and normal samples among the 3 data sets, respectively. The microarray data were first preprocessed using the algorithm "RMA", which contains background adjustment and normalization with quantile method. Then, the Moderated T statistic approach was used to select significant DEGs with the "limma" package of Bioconductor. Finally, DEGs were annotated through an annotation table downloaded from the GEO website. The *P* value <0.05 and |logFC| >1 were set as the cutoff criteria. The common DEGs from the 3 gene expression datasets were extracted via the online tool Venny 2.1 (*http://bioinfogp.cnb.csic.es/tools/venny/index.html*) and used for the following analyses. The heat maps of the common DEGs for each data set were constructed using the Morpheus online tool (*https://software.broadinstitute.org/morpheus/*).

Functional and pathway enrichment analysis

Gene ontology (GO, *http://www.geneontology.org/*) is a framework for the model of biology, which describes gene functions and classifies them along 3 aspects: molecular function (MF), cellular component (CC), and biological process (BP). The Kyoto Encyclopedia of Genes and Genomes (KEGG, *http://www.genome.jp/kegg/*) is a collection of databases for understanding high-level gene functions and utilities of the biological system. The Database for Annotation, Visualization, and Integrated Discovery (DAVID, *https://david.ncifcrf.gov/home.jsp*) is an online tool that provides a comprehensive set of functional annotation tools to investigate the biological meaning behind a mass of genes. In order to analyze DEGs at the functional level, GO term enrichment analysis and KEGG pathway enrichment analysis were conducted using DAVID. P<0.05 was considered to have statistical significance and to achieve significant enrichment.

PPI network construction and modules analysis

The Search Tool for the Retrieval of Interacting Genes (STRING, *https://string-db.org/*) is a freely accessible biological database to evaluate protein–protein interaction information. STRING (version 10) contains 9 643 763 proteins from 2031 organisms.
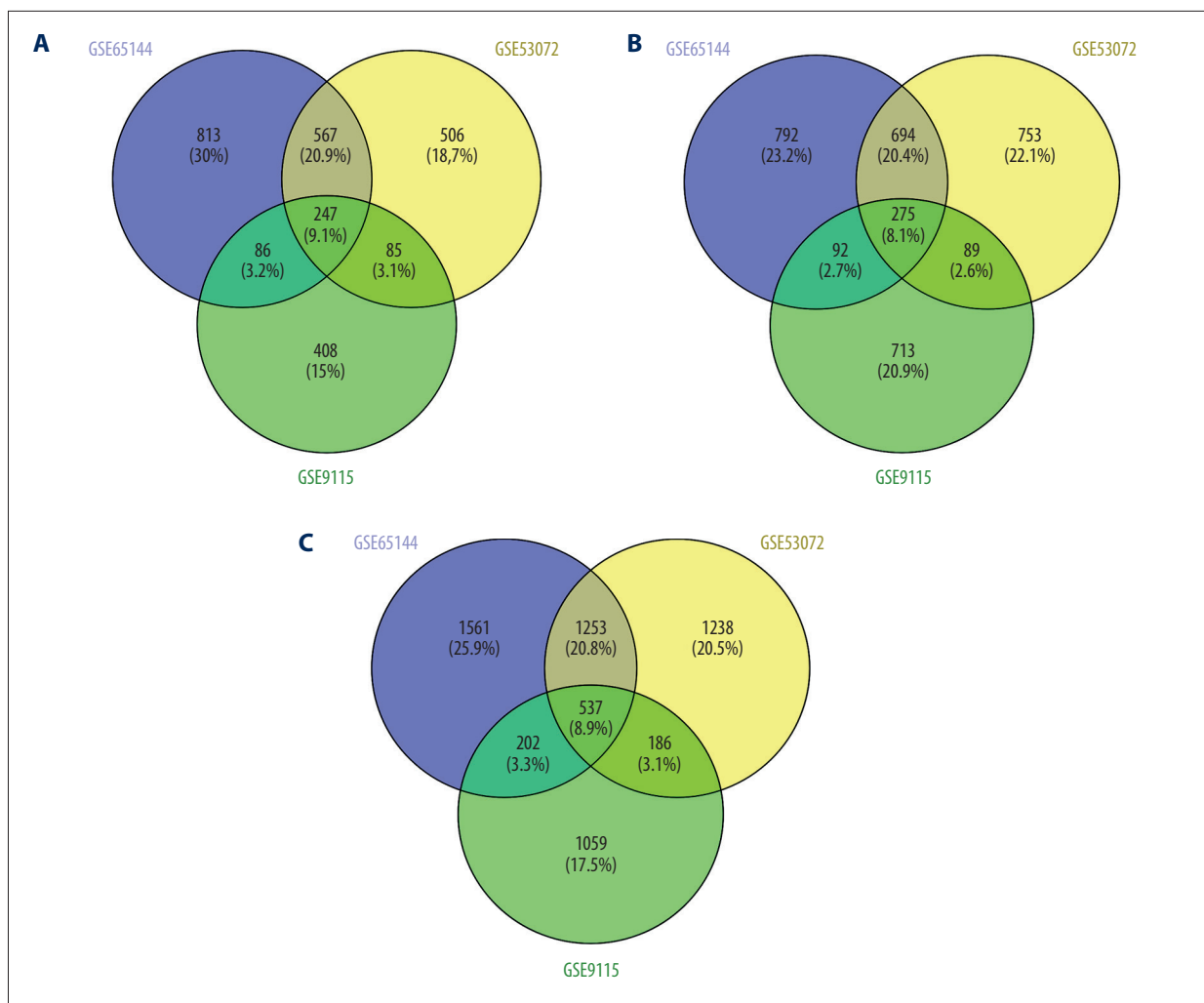
Figure 1. Identification results of differentially expressed genes in gene expression profiling datasets GSE53072, GSE65144, and GSE9115. (**A**) Upregulated DEGs of 3 datasets, (**B**) Downregulated DEGs of 3 datasets, (**C**) DEGs of the 3 data sets.

To assess the interactive relationships among DEGs, the DEGs were mapped to STRING with the confidence score >0.7 set as the cutoff criterion. Then, the PPI network was constructed and visualized using Cytoscape. The plug-in Molecular Complex Detection (MCODE) [16] was used to screen the modules of the PPI network with the following criteria: Max. Depth=100, K-Core=2, Mode Score Cutoff=0.2, and Degree Cutoff=2. Moreover, the function and pathway enrichment analysis were conducted for DEGs in each module with the plug-in Biological Network Gene Ontology (BiNGO). P<0.05 was considered to indicate significant differences.

Results

Identification of DEGs

A total of 3214, 3553, and 1884 DEGs were identified from GSE53072, GSE65144, and GSE9115, respectively, and a total

of 537 DEGs were screened out from all 3 datasets as shown in Figure 1C. Among them, 247 genes were upregulated and 275 genes were downregulated (Figure 1A, 1B). The hierarchical clustering analysis of 537 overlapped DEGs in the 3 datasets is shown in Figure 2. As shown in the expression heat maps, there were significant differences between the ATC group and the normal control group in expression patterns of DEGs. The results indicated the high consistency in expression patterns of DEGs among the 3 datasets, and that the overlapped DEGs may be truly differentially expressed in the ATC group compared with those in the normal control group.

GO term enrichment analysis

The common identified DEGs, including both upregulated and downregulated ones, were all uploaded to DAVID to identify overrepresented GO categories and KEGG pathways. The upregulated DEGs were mainly involved in biological process (BP),
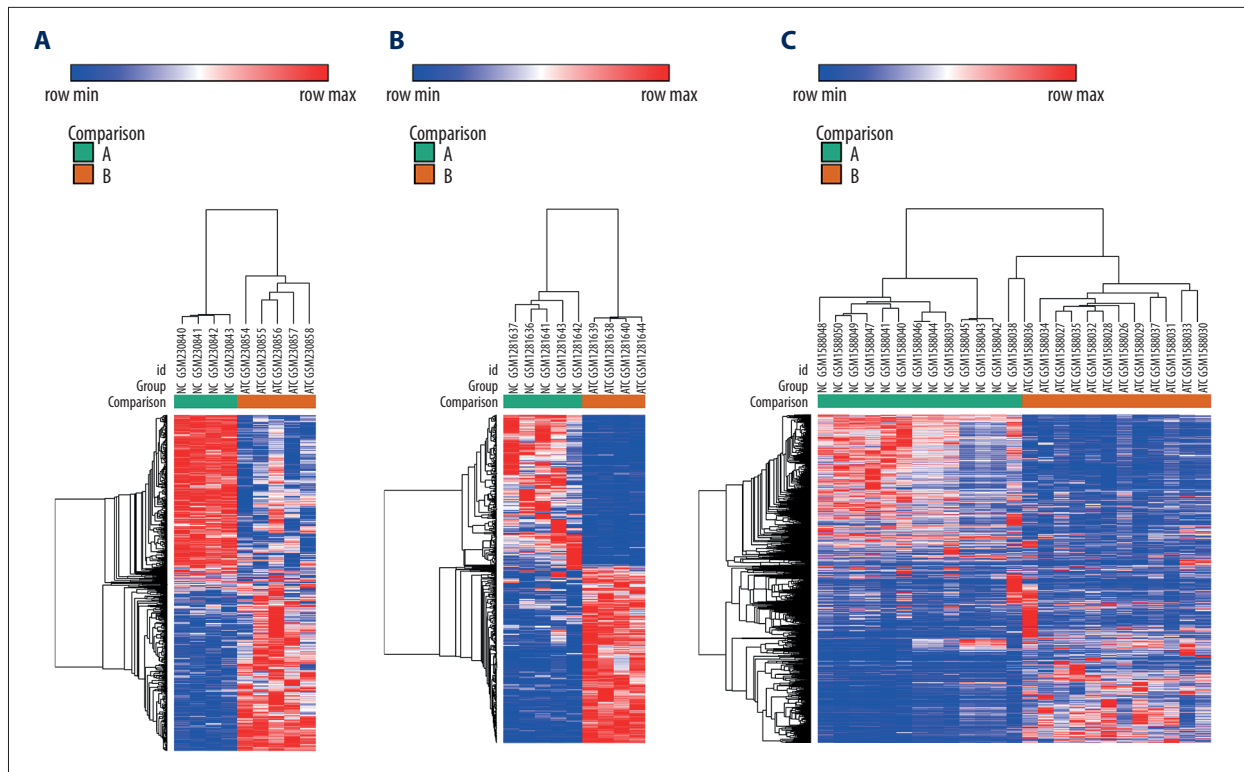
Figure 2. Heat map plots of differentially expressed genes across 3 gene expression profiles. (**A–C**) Heat maps of overlapped DEGs observed in datasets GSE9115, GSE53072 and GSE 65144, respectively. *NC* – normal control, *ATC* – anaplastic thyroid carcinoma.

associated with cell division, mitotic nuclear division, and cell adhesion, and the downregulated DEGs were significantly enriched in melanocyte differentiation, regulation of phosphatidylinositol 3-kinase signaling, and ventricular cardiac muscle cell action potential (Table 1). In the cell component (CC) ontology, the upregulated DEGs were enriched in kinetochore, extracellular space, and membrane, and the downregulated DEGs were enriched in extracellular exosome, apical plasma membrane, and sarcolemma. Molecular function (MF) analysis also indicated that the upregulated DEGs were significantly enriched in protein binding, ATP binding, and integrin binding, and the downregulated DEGs were enriched in calmodulin binding, oxidoreductase activity, and acyl-CoA dehydrogenase activity.

KEGG pathway analysis

To gain further insight into the function of the identified DEGs, Table 2 shows the top 10 significantly enriched pathways of the upregulated DEGs and downregulated DEGs through KEGG analysis. The upregulated DEGs were mainly enriched in cell cycle, ECM-receptor interaction, and leishmaniasis, and the downregulated DEGs were mainly enriched in thyroid hormone synthesis, insulin resistance, and pathways in cancer.

PPI network construction and module selection

Based on 537 DEGs, |logFC|>2 is individually set as the cutoff criteria to present a clear PPI network (Figure 3). Based on information from the STRING database, the top 10 DEGs with higher degrees were screened and selected as hub genes, which showed a strong association with other node proteins, as depicted in Figure 4. Among these hub genes, CDK1 showed the highest node degree, which was 113. In addition, the top 2 most significant modules were selected (Figure 5), and the enrichment analysis (Tables 3, 4) showed that the genes in these 2 modules were mainly associated with cell cycle and positive regulation of mitosis, respectively.

Discussion

Anaplastic thyroid cancer (ATC) remains one of the deadest diseases. Recent studies showed that mutations, amplifications, activation of oncogenes, and silencing of tumor suppressor genes contribute to its aggressive behavior [17]. Understanding the etiological factors and molecular mechanism of ATC is of critical importance to improving its survival rate and prevention. Due to the rapid development of microarray technologies in recent years, we are able to explore the potential targets for

Table 1. Gene ontology analysis of differentially expressed genes associated with anaplastic thyroid cancer.

Category	Term	Count	%	P value
Upregulated				
GOTERM_BP_DIRECT	GO: 0051301~cell division	29	0.073774454	8.89E-14
GOTERM_BP_DIRECT	GO: 0007067~mitotic nuclear division	20	0.050878934	2.14E-09
GOTERM_BP_DIRECT	GO: 0007155~cell adhesion	24	0.06105472	1.32E-07
GOTERM_BP_DIRECT	GO: 0007229~integrin-mediated signaling pathway	12	0.03052736	1.36E-07
GOTERM_BP_DIRECT	GO: 0000082~G1/S transition of mitotic cell cycle	12	0.03052736	1.86E-07
GOTERM_CC_DIRECT	GO: 0000776~kinetochore	12	0.03052736	9.65E-09
GOTERM_CC_DIRECT	GO: 0016020~membrane	56	0.142461014	2.23E-06
GOTERM_CC_DIRECT	GO: 0005615~extracellular space	40	0.101757867	3.20E-06
GOTERM_CC_DIRECT	GO: 0005925~focal adhesion	18	0.04579104	1.99E-05
GOTERM_CC_DIRECT	GO: 0000777~condensed chromosome kinetochore	9	0.02289552	2.04E-05
GOTERM_MF_DIRECT	GO: 0005515~protein binding	160	0.407031469	2.37E-06
GOTERM_MF_DIRECT	GO: 0005524~ATP binding	42	0.106845761	2.64E-05
GOTERM_MF_DIRECT	GO: 0005178~integrin binding	9	0.02289552	1.19E-04
GOTERM_MF_DIRECT	GO: 0005518~collagen binding	7	0.017807627	1.95E-04
GOTERM_MF_DIRECT	GO: 0004872~receptor activity	12	0.03052736	2.65E-04
Downregulated				
GOTERM_BP_DIRECT	GO: 0086005~ventricular cardiac muscle cell action potential	4	0.011480265	2.30E-04
GOTERM_BP_DIRECT	GO: 0014066~regulation of phosphatidylinositol 3-kinase signaling	7	0.016072372	0.00136878
GOTERM_BP_DIRECT	GO: 0030318~melanocyte differentiation	5	0.009184212	0.001486948
GOTERM_BP_DIRECT	GO: 0033539~fatty acid beta-oxidation using acyl-CoA dehydrogenase	4	0.009184212	0.002576085
GOTERM_BP_DIRECT	GO: 0009083~branched-chain amino acid catabolic process	4	0.009184212	0.003024104
GOTERM_CC_DIRECT	GO: 0070062~extracellular exosome	77	0.176796088	9.63E-08
GOTERM_CC_DIRECT	GO: 0016324~apical plasma membrane	16	0.036736849	3.52E-05
GOTERM_CC_DIRECT	GO: 0042383~sarcolemma	8	0.018368425	2.86E-04
GOTERM_CC_DIRECT	GO: 0016323~basolateral plasma membrane	11	0.025256584	3.96E-04
GOTERM_CC_DIRECT	GO: 0005856~cytoskeleton	16	0.036736849	5.10E-04
GOTERM_MF_DIRECT	GO: 0005516~calmodulin binding	12	0.027552637	1.40E-04
GOTERM_MF_DIRECT	GO: 0052890~oxidoreductase activity, acting on the CH-CH group of donors, with a flavin as acceptor	4	0.009184212	0.001339754
GOTERM_MF_DIRECT	GO: 0005200~structural constituent of cytoskeleton	8	0.018368425	0.001372584
GOTERM_MF_DIRECT	GO: 0003995~acyl-CoA dehydrogenase activity	4	0.009184212	0.001630695
GOTERM_MF_DIRECT	GO: 0044325~ion channel binding	7	0.016072372	0.007246827

Table 2. The top 10 enriched KEGG pathways of differentially expressed genes associated with anaplastic thyroid cancer.

| Pathway ID/Name	Count	%	P Value	Genes
Upregulated DEGs				
hsa04110: Cell cycle	14	0.035615254	6.15E-07	CDC7, CDC6, MAD2L1, CCNB2, YWHAH, BUB1, TTK, BUB1B, CHEK1, CCNA2, MCM4, MYC, MCM5, MCM6
hsa04512: ECM-receptor interaction	9	0.02289552	2.49E-04	COL3A1, ITGB5, ITGA2, ITGA4, ITGB1, COL5A1, SPP1, FN1, HMMR
hsa05140: Leishmaniasis	8	0.020351573	3.99E-04	IRAK1, MYD88, TLR2, FCGR2A, ITGA4, STAT1, ITGB1, ITGAM
hsa05205: Proteoglycans in cancer	12	0.03052736	0.001565479	CASP3, HPSE, MMP9, TLR2, ITGB5, ITGA2, MYC, ITGB1, PLAU, FN1, TWIST1, PLAUR
hsa04151: PI3K-Akt signaling pathway	16	0.040703147	0.002476592	IL2RA, COL3A1, TLR2, ITGB5, ITGA2, ITGA4, ITGB1, COL5A1, EIF4EBP1, YWHAH, IL4R, MYC, SPP1, FN1, CSF1R, F2R
hsa04145: Phagosome	10	0.025439467	0.00265876	CYBB, FCGR2B, TLR2, ITGB5, ITGA2, FCGR2A, CTSS, TUBA1A, ITGB1, ITGAM
hsa04611: Platelet activation	9	0.02289552	0.003464929	ADCY7, COL3A1, ITGA2, FCGR2A, ITGB1, VASP, COL5A1, LCP2, F2R
hsa04380: Osteoclast differentiation	9	0.02289552	0.003631749	CYBB, FCGR2B, LILRB4, FCGR2A, TREM2, STAT1, CSF1R, LCP2, TYROBP
hsa03030: DNA replication	5	0.012719733	0.004794752	POLE2, MCM4, MCM5, FEN1, MCM6
hsa04666: Fc gamma R-mediated phagocytosis	7	0.017807627	0.00540983	ARPC1B, FCGR2B, RAC2, HCK, WASF1, FCGR2A, VASP
Downregulated DEGs				
hsa04918: Thyroid hormone synthesis	7	0.016072372	0.001859153	SLC26A4, PLCB4, ADCY9, PAX8, ATP1A1, LRP2, ITPR1
hsa05202: Transcriptional misregulation in cancer	9	0.020664478	0.013054046	HHEX, NUPR1, ID2, CCND2, PAX8, ARNT2, TSPAN7, PBX1, MLLT3
hsa04931: Insulin resistance	7	0.016072372	0.015302523	PRKCQ, PIK3CB, RPS6KA2, TBC1D4, IRS1, PPARGC1A, SLC27A2
hsa04015: Rap1 signaling pathway	10	0.022960531	0.016414205	FGFR2, PLCB4, MAGI1, ADCY9, PIK3CB, EFNA1, VEGFA, CDH1, KIT, RAPGEF3
hsa04972: Pancreatic secretion	6	0.013776319	0.02955826	PLCB4, ADCY9, RYR2, ATP1A1, SLC4A4, ITPR1
hsa05200: Pathways in cancer	14	0.032144743	0.031228341	FGFR2, PIK3CB, ARNT2, FZD3, CDH1, KIT, DAPK2, EDNRB, FOS, PLCB4, ADCY9, BCL2, PAX8, VEGFA
hsa04960: Aldosterone-regulated sodium reabsorption	4	0.009184212	0.035652255	PIK3CB, ATP1A1, SCNN1A, IRS1
hsa04020: Calcium signaling pathway	8	0.018368425	0.049708184	EDNRB, GNAL, PLCB4, ADCY9, ERBB3, PDE1A, RYR2, ITPR1
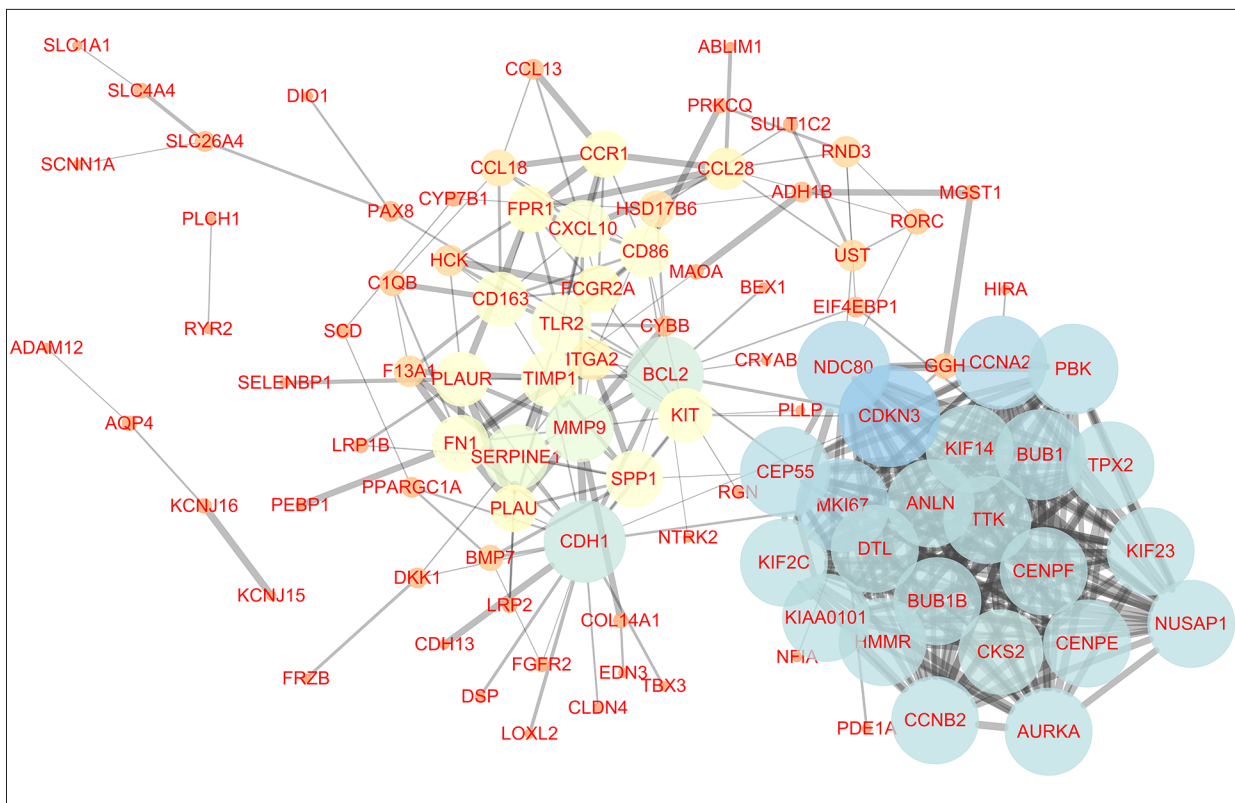
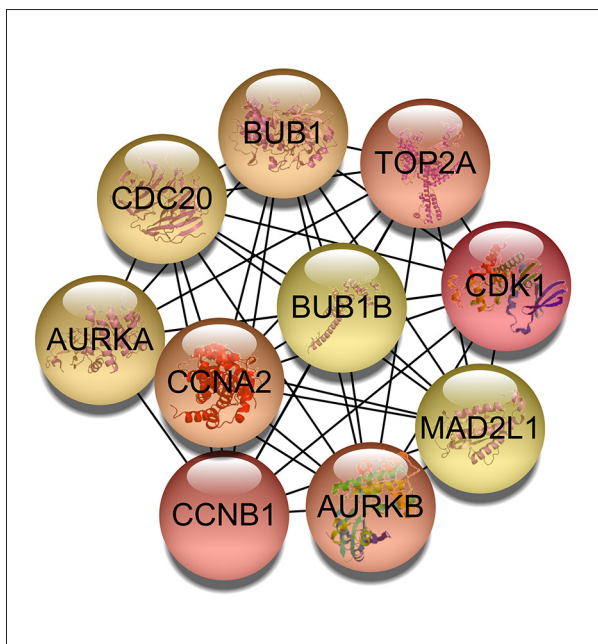**Figure 3.** Protein–protein interaction networks for all DEGs.

Figure 4. Top 10 hub genes in protein–protein interaction networks.

diagnosis, therapy, and prognosis of ATC via integrated bio-informatic methods.

In this study, 3 gene expression data sets of ATC were retrieved from the GEO data set. A total of 537 DEGs were screened, including 247 upregulated genes and 275 downregulated genes. In order to investigate the biological meaning behind these DEGs, we performed GO and KEGG pathway analysis.

The bioinformatics analysis showed that these upregulated genes were mainly related to cell proliferation. Specifically, GO term enrichment analysis showed that 29 upregulated DEGs participated in cell division and 20 upregulated DEGs enriched in mitotic nuclear division. KEGG pathway analysis indicated that 14 upregulated DEGs were significantly enriched in cell cycle and 5 upregulated DEGs were enriched in DNA replication. PPI network and module analysis also indicated that the gene modules were significantly enriched in cell proliferation. It has been shown that uncontrolled cell proliferation is the primary feature of cancer, and tumor cells can damage genes that directly regulate their cell cycles [18]. The role of the cell cycle is believed to be important in the etiology of cancer [19]. The present study clearly shows that many DEGs influenced the development of ATC by regulating cell proliferation. For ex-ample, CDCA4, a member of the CDCA family, which was up-regulated in our study, participates in regulation of prolifera-tion through the E2F/retinoblastoma pathway [20]. In addition, as the main upregulated DEGs, kinesin superfamily proteins (KIFs) are essential for cell mitosis and meiosis, the abnormality
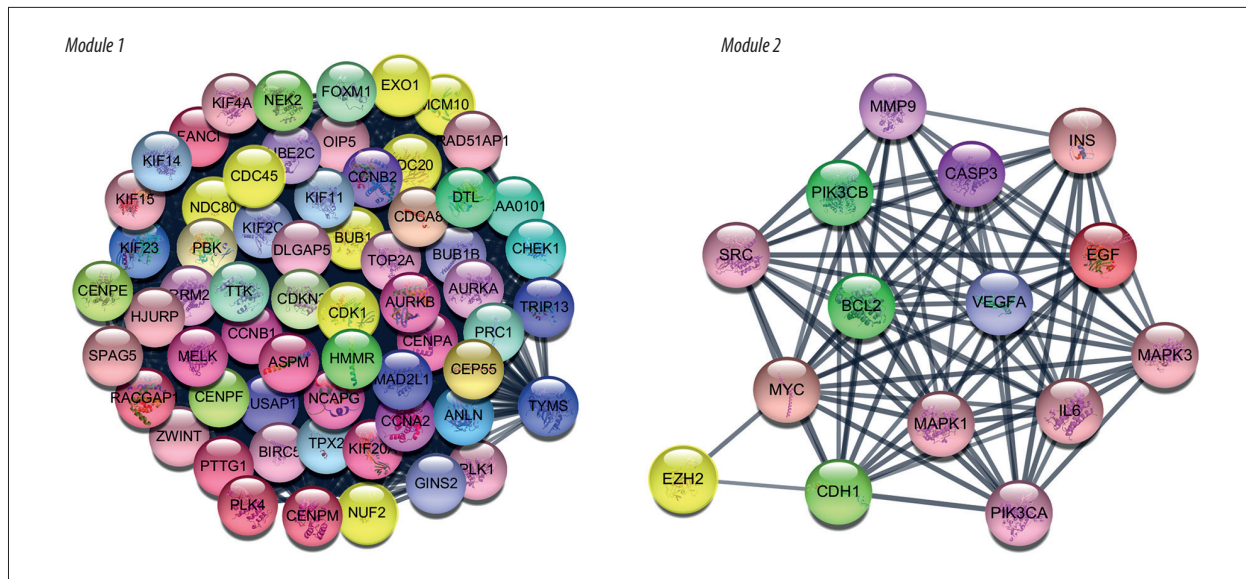
Figure 5. Top 2 significant modules from the protein–protein interaction network.

Table 3. GO enrichment analysis of module 1 from the protein–protein interaction network.

GO-ID	p Value	Description	Genes in test set
278	1.82E-08	Mitotic cell cycle	CDC20|CDC45|NUF2|CDK1|MCM10|BUB1|NDC80
22402	1.84E-08	Cell cycle process	CDC20|CDC45|EXO1|NUF2|CDK1|MCM10|BUB1|NDC80
7049	6.68E-08	Cell cycle	CDC20|CDC45|EXO1|NUF2|CDK1|MCM10|BUB1|NDC80
22403	1.76E-07	Cell cycle phase	CDC20|CDC45|NUF2|CDK1|MCM10|BUB1|NDC80
7067	2.29E-06	Mitosis	CDC20|NUF2|CDK1|BUB1|NDC80
87	2.41E-06	M phase of mitotic cell cycle	CDC20|NUF2|CDK1|BUB1|NDC80
280	2.98E-06	Nuclear division	CDC20|NUF2|CDK1|BUB1|NDC80
48285	3.84E-06	Organelle fission	CDC20|NUF2|CDK1|BUB1|NDC80
226	9.12E-06	Microtubule cytoskeleton organization	CDC20|NUF2|BUB1|NDC80
7017	1.60E-05	Microtubule-based process	CDC20|NUF2|BUB1|NDC80
45787	3.91E-05	Positive regulation of cell cycle	CDC20|CDK1

of which can cause cell death, gene deletion, and even carcinogenesis [21]. Therefore, it is meaningful to further explore the relationships and mechanisms between the DEGs and ATC.

GO term enrichment analysis also suggested that several downregulated genes participated in the development of thyroid carcinoma, as 4 genes were enriched in ventricular cardiac muscle cell action potential. Specifically, the first published GWAS in differentiated thyroid cancer confirmed the association with the 3p22 SNP rs1159444 c(GPD1L) [22]. In addition, the other 3 downregulated genes in ventricular cardiac muscle cell action potential have been confirmed to play roles in thyroid cancer [23,24], but the exact mechanism is still unknown.

Furthermore, KEGG pathway analysis demonstrated that downregulated DEGs are mainly involved in insulin resistance (IR) and pathways in cancer. This is consistent with data showing that high prevalence of IR is an important risk factor for thyroid cancer [25], and that the variant signaling pathway participates in thyroid cancer [26]. Hayashi et al. [27] reported the antitumor effects of peroxisome proliferator activate receptor γ ligands on anaplastic thyroid carcinoma, which has been regarded as a target for insulin resistance type 2 diabetes treatment. The enriched DEGs analyzed by the KEGG pathway are mainly involved in the PI3K-AKT signaling pathway, the MAPK signaling pathway, and the wnt signaling pathway. In previous studies, the PI3K-AKT and MAPK signaling pathways were shown

Table 4. GO enrichment analysis of module 2 from the protein–protein interaction network.

GO-ID	p Value	Description	Genes in test set
45840	3.22E-04	Positive regulation of mitosis	CDH1
45842	3.22E-04	Positive regulation of mitotic metaphase/anaphase transition	CDH1
51488	3.22E-04	Activation of anaphase-promoting complex activity	CDH1
45862	3.22E-04	Positive regulation of proteolysis	CDH1
51785	3.22E-04	Positive regulation of nuclear division	CDH1
32436	3.22E-04	Positive regulation of proteasomal ubiquitin-dependent protein catabolic process	CDH1
7092	3.22E-04	Activation of mitotic anaphase-promoting complex activity	CDH1
10697	3.22E-04	Negative regulation of spindle pole body separation	CDH1
51437	3.22E-04	Positive regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	CDH1
32434	4.83E-04	Regulation of proteasomal ubiquitin-dependent protein catabolic process	CDH1

to be primarily involved in FTC and differentiated PTC, respectively [28,29]. A growing number of studies have demonstrated that the dual activation of MAPK and PI3K-AKT pathways may be a common mechanism for promoting the progression of thyroid cancer [30]. Furthermore, Kurihara et al. have identified the association of the wnt signaling pathway with anaplastic thyroid cancer [31]. Hence, the molecular-targeting therapy of ATC needs to be further explored in future studies.

The PPI network was constructed among 569 DEGs and presented the top 10 hub genes: CDK1, CCNB1, TOP2A, AURKB, CCNA2, BUB1, AURKA, CDC20, MAD2L1, and BUB1B. CDK1 has the highest degree of node connectivity in the PPI network; it plays an important role in unrestricted cell proliferation and regulates the cell cycle, mainly by interacting with cell cycle regulatory cyclins. For example, the CDK1-cyclin B1 complex is responsible for entry and progression of the mitotic phase (M) [32,33]. By inhibiting the activity of CDK1, the proliferation of thyroid cancer can be repressed *in vivo* and *in vitro* [34]. The second and the fifth hub genes, CCNB1 and CCNA2, belong to the highly conserved cyclin family. CCNB1 complexes with p43 (cdc2) to form the maturation-promoting factor (MPF) and plays vital roles in control of the cell cycle at G2/M transitions, while CCNA2 is involved in control of the cell cycle at the G1/S and the G2/M transitions and is necessary in embryonic cells and hematopoietic lineage. Both of them have been reported to be expressed in many cancer types [35,36]. Accumulating evidence shows that CCNB1 is overexpressed in 4 undifferentiated thyroid carcinomas [37] and ATC is overexpressed in genes that code for cyclins, including CCNA2 [13]. The third hub gene TOP2A is located on chromosome 17q12-q21 near the HER2 oncogene. It is well known that co-amplification of HER2 and TOP2A is associated with sensitivity to

anthracycline therapy for some types of cancer [38]. Some recent studies also reported that TOP2A is upregulated in the ATC and cell lines and is involved in DNA replication and DNA metabolic processes [39]. As a member of the Aurora kinases family, AURKA is an important protein in the regulation of G2/M transition during mitosis. It has been identified that AURKA is one of the most frequently and strongly overexpressed genes in some tumors, including ATC [40]. Another Aurora kinase, AURKB, has a dynamic subcellular localization during cell cycle progression. It was shown that overexpression of AURKB can lead to multinucleation and polyploidy and is highly correlated with genomic instability and poor prognosis in thyroid cancers, indicating its role in neoplastic transformation [41].

BUB1 is a multi-task protein kinase required for proper chromosome segregation in eukaryotes. Quyang et al. [42] demonstrated that no mutations of BUB1 exist in the thyroid carcinoma cell lines and that other factors must account for the mitotic checkpoint dysfunction and chromosomal instability detected in thyroid cancer cells. In contrast, a recent study showed that the high expression of BUB1B, which originated from the same gene of BUB1, was associated with recurrence in papillary thyroid cancer [43]. Therefore, more studies are needed to investigate the function of BUB1 and BUB1B in ATC. The interaction between BUB1 and BUB1B in ATC needs to be explored as well. As an important factor of the spindle assembly checkpoint, CDC20 plays a crucial role in tumorigenesis and progression. Recent studies showed that CDC20 expression was ATC-specifically upregulated [44]. It has also been shown that disruption of MAD2L1, an important interactive protein in maintaining spindle checkpoint function, is closely associated with aneuploidy and tumorigenesis [45]. However, further studies are needed to determine the role of MAD2L1 in ATC.

Conclusions

In conclusion, we used a series of bioinformatics analysis methods to identify the key genes and pathways involved in ATC initiation and progression from 3 gene expression profiles containing normal samples and ATC samples. Our results provide a more detailed molecular mechanism for the development of ATC, shedding light on the potential biomarkers and therapeutic targets. However, the interacting mechanism and function of genes need to be confirmed in further experiments.

References:

1. Siegel R, DeSantis C, Virgo K et al: Cancer treatment and survivorship statistics, 2012. Cancer J Clin, 2012; 62(4): 220–41

2. Siegel R, Ma J, Zou Z, Jemal A: Cancer statistics, 2014. Cancer J Clin, 2014; 64(1): 9–29

3. Smallridge RC, Marlow LA, Copland JA: Anaplastic thyroid cancer: molecular pathogenesis and emerging therapies. Endocr Relat Cancer, 2009; 16(1): 17–44

4. Garg M, Kanojia D, Okamoto R et al: Laminin-5gamma-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, 2014; 99(1): E62–72

5. Kulasingam V, Diamandis EP: Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol, 2008; 5(10): 588–99

6. Wang Y, Zheng T: Screening of hub genes and pathways in colorectal cancer with microarray technology. Pathol Oncol Res, 2014; 20(3): 611–18

7. Wang J, Wei B, Cao S et al: Identification by microarray technology of key genes involved in the progression of carotid atherosclerotic plaque. Genes Genet Syst, 2014; 89 (6): 253–58

8. Liu YJ, Zhang S, Hou K et al: Analysis of key genes and pathways associated with colorectal cancer with microarray technology. Asian Pac J Cancer Prev, 02013; 14(3): 1819–23

9. Lin B, Madan A, Yoon JG et al: Massively parallel signature sequencing and bioinformatics analysis identifies up-regulation of TGFBI and SOX4 in human glioblastoma. PLoS One, 2010; 5(4): e10210

10. Goujon M, McWilliam H, Li W et al: A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res, 2010; 38(Web Server issue): W695–99

11. Rhodes DR, Yu J, Shanker K et al: ONCOMINE: A cancer microarray database and integrated data-mining platform. Neoplasia (New York, NY), 2004; 6(1): 1–6

12. Yu J, Mai W, Cui Y, Kong L: Key genes and pathways predicted in papillary thyroid carcinoma based on bioinformatics analysis. J Endocrinol Invest, 2016; 39(11): 1285–93

13. Salvatore G, Nappi TC, Salerno P et al: A cell proliferation and chromosomal instability signature in anaplastic thyroid carcinoma. Cancer Res, 2007; 67(21): 10148–58

14. von Roemeling CA, Marlow LA, Pinkerton AB et al: Aberrant lipid metabolism in anaplastic thyroid carcinoma reveals stearoyl CoA desaturase 1 as a novel therapeutic target. J Clin Endocrinol Metab, 2015; 100 (5): E697–709

15. Pita JM, Figueiredo IF, Moura MM et al: Cell cycle deregulation and TP53 and RAS mutations are major events in poorly differentiated and undifferentiated thyroid carcinomas. J Clin Endocrinol Metab, 2014; 99(3): E497–507

16. Bader GD, Hogue CW: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 2003; 4: 2

17. Smallridge RC, Copland JA: Anaplastic thyroid carcinoma: pathogenesis and emerging therapies. Clin Oncol, 2010; 22(6): 486–97

18. Sherr CJ: Cancer cell cycles. Science, 1996; 274(5293): 1672–77

19. Hunter T, Pines J: Cyclins and cancer. II: Cyclin D and CDK inhibitors come of age. Cell, 1994; 79(4): 573–82

20. Hayashi R, Goto Y, Ikeda R et al: CDCA4 is an E2F transcription factor family-induced nuclear factor that regulates E2F-dependent transcriptional activation and cell proliferation. J Biol Chem, 2006; 281(47): 35633–48

21. Yu Y, Feng YM: The role of kinesin family proteins in tumorigenesis and progression: potential biomarkers and molecular targets for cancer therapy. Cancer, 2010; 116(22): 5150–60

22. Gudmundsson J, Sulem P, Gudbjartsson DF et al: Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. Nat Genet, 2009; 41(4): 460–64

23. Kim HS, Kim DH, Kim JY et al: Microarray analysis of papillary thyroid cancers in Korean. Korean J Intern Med, 2010; 25(4): 399–407

24. Arora N, Scognamiglio T, Lubitz CC et al: Identification of borderline thyroid tumors by gene expression array analysis. Cancer, 2009; 115(23): 5421–31

25. Rezzonico JN, Rezzonico M, Pusiol E et al: Increased prevalence of insulin resistance in patients with differentiated thyroid carcinoma. Metabol Syndr Relat Disord, 2009; 7(4): 375–80

26. Xie J, Fan Y, Zhang X: Molecular mechanisms in differentiated thyroid cancer. Front Biosci (Landmark Ed) 2016; 21: 119–29

27. Hayashi N, Nakamori S, Hiraoka N et al: Antitumor effects of peroxisome proliferator activate receptor gamma ligands on anaplastic thyroid carcinoma. Int J Oncol, 2004; 24(1): 89–95

28. Wang Y, Hou P, Yu H et al: High prevalence and mutual exclusivity of genetic alterations in the phosphatidylinositol-3-kinase/akt pathway in thyroid tumors. J Clin Endocrinol Metab, 2007; 92(6): 2387–90

29. Santarpia L, El-Naggar AK, Cote GJ et al: Phosphatidylinositol 3-kinase/akt and ras/raf-mitogen-activated protein kinase pathway mutations in anaplastic thyroid cancer. J Clin Endocrinol Metab, 2008; 93(1): 278–84

30. Miller KA, Yeager N, Baker K et al: Oncogenic Kras requires simultaneous PI3K signaling to induce ERK activation and transform thyroid epithelial cells in vivo. Cancer Res, 2009; 69(8): 3689–94

31. Kurihara T, Ikeda S, Ishizaki Y et al: Immunohistochemical and sequencing analyses of the Wnt signaling components in Japanese anaplastic thyroid cancers. Thyroid, 2004; 14(12): 1020–29

32. Malumbres M, Barbacid M: Cell cycle, CDKs and cancer: A changing paradigm. Nat Rev Cancer, 2009; 9(3): 153–66

33. Merrick KA, Fisher RP: Why minimal is not optimal: driving the mammalian cell cycle – and drug discovery – with a physiologic CDK control network. Cell Cycle (Georgetown, Tex), 2012; 11(14): 2600–5

34. Lin SF, Lin JD, Hsueh C et al: A cyclin-dependent kinase inhibitor, dinaciclib in preclinical treatment models of thyroid cancer. PLoS One, 2017; 12(2): e0172315

35. Hassan KA, Ang KK, El-Naggar AK et al: Cyclin B1 overexpression and resistance to radiotherapy in head and neck squamous cell carcinoma. Cancer Res, 2002; 62(22): 6414–17

36. Uhlen M, Oksvold P, Fagerberg L et al: Towards a knowledge-based Human Protein Atlas. Nat Biotechnol, 2010; 28(12): 1248–50

37. Ito Y, Yoshida H, Nakano K et al: Expression of G2-M modulators in thyroid neoplasms: Correlation of cyclin A, B1 and cdc2 with differentiation. Pathol Res Pract, 2002; 198(6): 397–402

38. Liang Z, Zeng X, Gao J et al: Analysis of EGFR, HER2, and TOP2A gene status and chromosomal polysomy in gastric adenocarcinoma from Chinese patients. BMC Cancer, 2008; 8: 363

39. Saiselet M, Floor S, Tarabichi M et al: Thyroid cancer cell lines: an overview. Front Endocrinol (Lausanne), 2012; 3: 133

40. Baldini E, Sorrenti S, D'Armiento E et al: Aurora kinases: New molecular targets in thyroid cancer therapy. Clin Ter, 2012; 163 (6): e457–62

41. Tsou JH, Chang KC, Chang-Liao PY et al: Aberrantly expressed AURKC enhances the transformation and tumourigenicity of epithelial cells. J Pathol, 2011; 225(2): 243–54

42. Ouyang B, Knauf JA, Ain K et al: Mechanisms of aneuploidy in thyroid cancer cell lines and tissues: evidence for mitotic checkpoint dysfunction without mutations in BUB1 and BUBR1. Clin Endocrinol, 2002; 56(3): 341–50

43. Chien MN, Yang PS, Lee JJ et al: Recurrence-associated genes in papillary thyroid cancer: An analysis of data from The Cancer Genome Atlas. Surgery, 2017; 161(6): 1642–50

44. Gayyed MF, El-Maqsoud NM, Tawfiek ER et al: A comprehensive analysis of CDC20 overexpression in common malignant tumors from multiple organs: Its correlation with tumor grade and stage. Tumour Biol, 2016; 37(1): 749–62

45. Guo Y, Zhang X, Yang M et al: Functional evaluation of missense variations in the human MAD1L1 and MAD2L1 genes and their impact on susceptibility to lung cancer. J Med Genet, 2010; 47(9): 616–22