Investigating the mechanisms of papillary thyroid carcinoma using transcriptome analysis

JIE QIU1*, WENWEI ZHANG2*, QINGSHENG XIA3, FUXUE LIU4, SHUWEI ZHAO5, KAILING ZHANG1, MIN CHEN1, CHUANSHAN ZANG1, RUIFENG GE1, DAPENG LIANG1 and YAN SUN1

1Otolaryngology Head and Neck Surgery; 2Radiology Department, The Affiliated Hospital of Qingdao University, Qingdao, Shandong 266003; 3Otolaryngology Head and Neck Surgery, Qingdao Municipal Hospital, Qingdao, Shandong 266071;
4Otolaryngology Head and Neck Surgery, Shaoxing Municipal Hospital, Shaoxing, Zhejiang 312000;
5Otolaryngology Head and Neck Surgery, Shanghai Chang Zheng Hospital, Shanghai 200003, P.R. China

Received October 13, 2016; Accepted June 12, 2017
DOI: 10.3892/mmr.2017.7346

Abstract. As the predominant thyroid cancer, papillary thyroid cancer (PTC) accounts for 75-85% of thyroid cancer cases. This research aimed to investigate transcriptomic changes and key genes in PTC. Using RNA-sequencing technology, the transcriptional profiles of 5 thyroid tumor tissues and 5 adjacent normal tissues were obtained. The single nucleotide polymorphisms (SNPs) were identified by SAMtools software and then annotated by ANNOVAR software. After differentially expressed genes (DEGs) were selected by edgR software, they were further investigated by enrichment analysis, protein domain analysis, and protein-protein interaction (PPI) network analysis. Additionally, the potential gene fusion events were predicted using FusionMap software. A total of 70,172 SNPs and 2,686 DEGs in the tumor tissues, as well as 83,869 SNPs in the normal tissues were identified. In the PPI network, fibronectin 1 (FN1; degree=31) and transforming growth factor β receptor 1 (TGFβR1; degree=22) had higher degrees. A total of 7 PPI pairs containing the non-synonymous risk SNP loci might serve a role in PTC by interacting with each other.

Introduction

Thyroid cancer originates from parafollicular or follicular thyroid cells, and includes papillary thyroid cancer (PTC), anaplastic thyroid cancer, medullary thyroid cancer, poorly differentiated thyroid cancer and follicular thyroid cancer (1). PTC accounts for 75-85% of all thyroid cancer cases and thus is the predominant thyroid cancer (2). It often occurs in young females, and is the most common thyroid cancer in children and patients who have undergone radiation therapy to the head and neck (3).

Research has focused on the pathogenesis of PTC. For instance, point mutation of serine-threonine protein kinase B-RAF (BRAF) occurs in approximately one-third to one-half of PTC cases, and BRAF can result in the activation of the carcinogenic mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase signaling pathway (4). The BRAF mutation in PTC patients is associated with poorer clinicopathological outcomes and can be used to predict recurrence independently; therefore, the BRAF mutation may serve as a promising marker for risk assessment of PTC (5-7). Stephens et al (8) investigated the loss of heterozygosity for three single nucleotide polymorphisms (SNPs; G691S, S904S, and L769L) of ret proto-oncogene (RET) in thyroid tumor and normal tissues, and demonstrated that RET SNPs may function in sporadic PTC. Salvatore et al (9) demonstrated that both RET/PTC rearrangements and the BRAF mutation are markers of PTC and can be utilized for fine-needle aspiration in conjunction with traditional cytology. Vascular endothelial growth factor regulates cancer-associated neo-angiogenesis and progression, and its expression and polymorphisms may indicate the aggressiveness behavior of PTC (10). The A-kinase anchoring protein 9-BRAF gene fusion, which is induced by the BRAF rearrangement through paracentric inversion of chromosome 7q, serves a role in activating the MAPK pathway in thyroid cancer (11). However, the molecular mechanisms of PTC have not yet been fully investigated.
RNA-sequencing (RNA-seq), which is a useful tool for transcriptome analysis, can be applied to reveal genomic structural variations, gene fusion events, novel genes and transcripts (12,13). By RNA-seq, Costa et al (14) identified new missense mutations in Casitas B-cell lymphoma gene; NOTCH1; phosphoinositide-3-kinase regulatory subunit 4 and SW/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, member 4 genes; somatic mutations in dicer 1, ribonuclease type III; met proto-oncogene (MET); and von Hippel-Lindau genes; and a new chimeric transcript induced by the WNK lysine deficient protein kinase 1-ß, 4-N-acetylgalactosaminiltransferase-3 gene fusion in PTC patients. Smallridge et al (15) used RNA-seq to analyze differentially expressed genes (DEGs) between BRAF wild-type and BRAF V600E mutation PTCs, demonstrating that the BRAF V600E mutation inhibits expression of immune/inflammatory response genes but promotes expression of chemokine (C-X-C motif) ligand 14 and human leukocyte antigen-G. However, the above studies did not perform comprehensive bioinformatics analysis. In the present study, the transcriptional profiles of thyroid tumor tissues and adjacent normal tissues were obtained by RNA-seq. Thereafter, SNPs were identified and functionally annotated. In addition, the DEGs were screened and further investigated by enrichment analysis, protein domain analysis and protein-protein interaction (PPI) network analysis. Furthermore, the potential gene fusion events were separately predicted for the tumor tissues and normal tissues. The flow chart of the bioinformatics analysis is presented in Fig. 1.

Materials and methods

Sample source and RNA sample preparation. Thyroid tumor tissues and adjacent normal tissues were collected from 5 PTC patients [2 women and 3 men; mean age, 37.4 years; 3 patients in stage I (T1N1aM0) and 2 patients in stage III (T1N1aM0)] from The Affiliated Hospital of Qingdao University (Qingdao, China). All patients were treated by thyroidectomy and lymphadenectomy; ultrasonography confirmed that the patients had no obvious abnormality after surgery. Total RNA was isolated from the tumor and normal tissues, using a TRIzol total RNA extraction kit (Invitrogen; Thermofisher Scientific, Inc., Waltham, MA, USA). The concentration of RNA was tested by spectrophotometry using NanoDrop™ RNA Assay kit in Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Inc.). Next, the RNA integrity was assessed by spectrophotometry using NanoDrop™ 2000 (Thermo Fisher Scientific, Inc., Wilmington, DE, USA) at 260 and 280 nm and 1% (w/v) agarose gel electrophoresis at 37°C for 15 min with 95°C preheat for 5 min. The results were stained with ethidium bromide (0.50 µg/ml; Sigma-Aldrich; Merck KGaA, Darmstadt, Germany) and eluted with elution buffer, followed by end repair and the ligation of sequencing adaptors. After the fragments with suitable size were selected by agarose gel electrophoresis, polymerase chain reaction (PCR) amplification was performed. Finally, the cDNA libraries were constructed and sequenced on an Illumina HiSeq 2500 platform according the manufacturer's protocol.

SNP identification and DEG screening. The clean reads were mapped to the hg19 human reference genome by the Burrows-Wheeler Aligner (bio-bwa.sourceforge.net/bwa.shtml) software (16), and then SNPs were identified by the SAMtools software (samtools.sourceforge.net) (17). Using ANNOVAR software (www.openbioinformatics.org/annovar/) (18), functional annotation was performed for the identified SNPs to investigate their genomic locations and variation information.

The DEGs in the thyroid tumor tissues compared with the adjacent normal tissues were screened using edgR software (bioconductor.org/packages/2.4/bioc/html/edgeR.html) (19). Based on the negative binomial model (20), the gene significance levels between the two groups were calculated. Following this, the adjusted P-value (the false discovery rate, FDR) of each gene was calculated by the Benjamini-Hochberg method (21). FDR<0.05 and |log2 fold-change (FC)|≥1 were considered thresholds.

Gene fusion analysis and enrichment analysis. The potential gene fusion events were predicted using FusionMap software (www.omicsoft.com/fusionmap/) (22), and then visualized using Circos software (www.circos.ca/) (23). The Gene Ontology (GO; www.geneontology.org/) database describes cellular component (CC), molecular function (MF), and biological process (BP) of gene products (24). The Kyoto Encyclopedia of Genes and Genomes (KEGG; www.genome.ad.jp/kegg/) database, which consists of known genes and their functions, can be used for pathway mapping (25). Using the Database for Annotation, Visualization and Integrated Discovery tool (DAVID; version 6.8; david.abcc.ncifcrf.gov/) software (26), GO functional and KEGG pathway enrichment analyses were conducted for the identified DEGs. P<0.05 was considered as the cut-off criterion.
PPI network analysis and protein domain analysis. The interactions among the proteins encoded by the DEGs were predicted using the Human Protein Reference Database (www.hprd.org) (27), and then PPI network was visualized using Cytoscape software (version 3.2.0; www.cytoscape.org/) (28). Using pfam_scan tool (www.ebi.ac.uk/Tools/pfa/pfam-scan/) (29), the domains in the proteins containing risk SNP loci were predicted, and the domains in the proteins containing non-synonymous risk SNP loci were selected. In addition, the non-synonymous risk SNP loci in the interaction domains involved in the PPI pairs were identified using the Database of Protein Domain Interactions (domine.utdallas.edu/cgi-bin/Domine) (30).

Results

SNP analysis. Upon mapping the clean reads to human reference genome hg19, 70,172 and 83,869 SNPs separately were identified in all of the thyroid tumor tissues and all of the adjacent normal tissues. A Venn diagram demonstrated that 18,795 SNPs were specific in the thyroid tumor tissues and were considered risk SNPs for PTC (Fig. 2). Subsequently, functional annotation was performed for these risk SNPs. Among the 18,795 SNPs, 739 risk SNPs were located in the exon regions (Fig. 3A). The variation information of the risk SNPs located in the exon regions is presented in Fig. 3B.

Integrated analysis of SNPs and DEGs. With the thresholds of FDR<0.05 and \( \log_{2} FC \geq 1 \), a total of 2,686 DEGs were screened in the thyroid tumor tissues compared with adjacent normal tissues, including 1,361 upregulated genes and 1,325 downregulated genes. The integrated analysis of SNPs and DEGs demonstrated that 12,528 risk SNPs were located in 4,317 genes (157 upregulated genes and 519 downregulated genes). The integrated analysis of SNPs and DEGs demonstrated that 12,528 risk SNPs were located in 4,317 genes (157 upregulated genes and 519 downregulated genes).

Using the pfam_scan tool, the domains in the proteins containing risk SNP loci or non-synonymous risk SNP loci were predicted. The results demonstrated that the risk SNPs had no significant effects on protein domains, whereas the non-synonymous risk SNPs may affect protein domains. The non-synonymous risk SNPs in the protein domains are listed in Table I.
Gene fusion analysis. The potential gene fusion events were predicted and are presented in Fig. 4. A total of 11 and 4 gene fusion events were identified in all of the thyroid tumor tissues and all of the normal tissues, respectively. A total of three gene fusion events were predicted in both thyroid tumor tissues and normal tissues, including RNA binding motif protein (RBM) 14-RBM 4, NK2 homeobox 1-surfactant associated 3 (NKX2-1-SFTA3), and chromosome 1 open reading frame 86 (C1orf86)-LOC100128003 gene fusions. The sodium channel, non-voltage gated 1 α-tumor necrosis factor receptor superfamily 1A gene fusion was predicted not in thyroid tumor tissues but in normal tissues. Furthermore, no gene fusion event was predicted in thyroid tumor tissues but not in normal tissues.

Enrichment analysis of the DEGs containing risk SNPs. The DEGs containing risk SNPs were investigated by functional and pathway enrichment analyses. The upregulated genes containing risk SNPs were significantly enriched in urogenital tissues.
system development (GO_BP, $P=1.38\times10^{-3}$), extracellular region part (GO_CC, $P=5.34\times10^{-3}$), hormone binding (GO_MF, $P=7.23\times10^{-3}$), and pathways in cancer (pathway, $P=1.28\times10^{-2}$) (Table IIA). Meanwhile, the significant terms enriched for the downregulated genes containing risk SNPs included regulation of cell morphogenesis (GO_BP, $P=4.41\times10^{-9}$), plasma membrane part (GO_CC, $P$-value = $2.22\times10^{-11}$), calcium ion binding (GO_MF, $P=6.18\times10^{-7}$) and cell adhesion molecules (pathway, $P=1.16\times10^{-6}$; Table IIB).

**PPI network analysis.** The PPI network constructed for the DEGs consisted of 716 nodes (273 upregulated genes and 443 downregulated genes) and 1,011 interactions (Fig. 5). Importantly, fibronectin 1 ($FN1$, degree=31) and transforming growth factor β ($TGF\beta$) receptor 1 ($TGF\beta R1$, degree=22) had higher degrees in the PPI network. Thereafter, the non-synonymous risk SNP loci in the interaction domains involved in the PPI pairs were identified. A total of 7 PPI pairs containing the non-synonymous risk SNP loci in the interaction domains were identified (Table III). Particularly, the interaction domains involved in the interactions of $FN1$ and 5 other proteins (such as tenascin C, $TNC$) contained non-synonymous risk SNP loci.

**Discussion**

There were some limitations in the present study (low case number, no own in vivo or in vitro experiments), thus the findings in five PTC cases had the characteristics of an advanced case report. In this study, a total of 70,172 and 83,869 SNPs were identified in all of the thyroid tumor tissues and all of the normal tissues, respectively. A Venn diagram demonstrated that 18,795 risk SNPs were specific in the thyroid tumor tissues. There were more SNPs in the normal tissues than in the thyroid tumor tissues; however, only the 18,795 SNPs specific in the thyroid tumor tissues were potential risk SNPs for PTC. Total 2,686 DEGs were screened in the thyroid tumor tissues, including 1,361 upregulated genes and 1,325 downregulated genes. The integrated analysis of SNPs and DEGs demonstrated that 12,528 risk SNPs were located in 4,317 genes (157 upregulated and 519 downregulated genes).

A previous study demonstrated that the activity of the $TGF\beta$/mothers against decapentaplegic (Smad)-dependent signaling pathway is associated with nodal metastasis, local invasion and $BRAF$-mutated PTCs (31). $TGF\beta 1$ has been identified as a key factor in PTC cells that affects the activation of stromal fibroblasts in a paracrine manner. Furthermore, the activation of the $TGF\beta$/Smad3 and Notch signaling pathways can impact tumor growth (32). Choe et al (33) investigated the association between SNPs (-1444C/G, Asn389Asn, -834 G/A) of $TGF\beta R2$ and PTC development, and the SNPs and clinico-pathological characteristics of PTC (including lymph node metastasis, location, size, number and extrathyroidal invasion), and demonstrated that $TGF\beta R2$ may serve a role in PTC progression in the Korean population. These studies suggested that $TGF\beta R$ had a correlation with the pathogenesis of PTC. In the present study, $FN1$ (degree=31) and $TGF\beta R1$ (degree=22) had higher degrees in the PPI network. Thus, $TGF\beta R1$ may serve an important role in the progression of PTC.
Table II. Significant GO and KEGG terms separately enriched for the upregulated and downregulated genes containing risk SNPs.

### A. Significant GO and KEGG terms enriched for the upregulated genes containing risk SNPs

| Category          | Term                                           | Gene number | Gene symbol                      | P-value      |
|-------------------|------------------------------------------------|-------------|----------------------------------|--------------|
| GO_BP             | GO:0001655~urogenital system development       | 6           | ACE, AR, BMP2, PGF, WFS1, ZBTB16  | 1.38E-03     |
| GO_BP             | GO:0055114~oxidation reduction                 | 13          | ALDH6A1, SORD, FOXRED2, OGDHL, AASS, CYP20A1, SOD3, IYD, ALDH1A1, ERO1LB, DIO2, BC02, WWOX | 2.83E-03     |
| GO_BP             | GO:0031667~response to nutrient levels         | 7           | BMP2, GATM, MAP1LC3A, PPAR, LIPG, ADIPOR2, GHR | 3.56E-03     |
| GO_BP             | GO:0001822~kidney development                  | 5           | ACE, BMP2, PGF, WFS1, ZBTB16      | 5.75E-03     |
| GO_BP             | GO:0043434~response to peptide hormone stimulus| 6           | AR, EGR2, GATM, PPAR, FOXO1, GHR  | 5.91E-03     |
| GO_CC             | GO:0044421~extracellular region part           | 18          | TG, MATN2, BMP2, PODN, SORD, PGF, MMP15, CCL28, SOD3, TNFSF10, ACE, FBLN2, EMID1, NUCB2, LIPG, ANGPTL1, MFA4, GHR, SOD3, CSGALNACT1, API53, SYNE1, EMID1, NUCB2, NDRG2, GRB14, WWOX, SMDP3 | 5.34E-03     |
| GO_CC             | GO:0005794~Golgi apparatus                     | 16          | TG, SLC39A14, RAP1GAP, NCS1, ACACB, MAN1A1, SOD3, CSGALNACT1, API53, SYNE1, EMID1, NUCB2, NDRG2, GRB14, WWOX, SMDP3 | 1.14E-02     |
| GO_CC             | GO:0005925~focal adhesion                      | 5           | TNS3, PGMS, PRUNE, LAYN, LPP      | 1.35E-02     |
| GO_CC             | GO:0005924~cell-substrate adherens junction    | 5           | TNS3, PGMS, PRUNE, LAYN, LPP      | 1.53E-02     |
| GO_CC             | GO:0031985~Golgi cisterna                      | 3           | CSGALNACT1, NUCB2, SMDP3          | 1.65E-02     |
| GO_MF             | GO:0042562~hormone binding                     | 4           | ALDH1A1, AR, ADIPOR2, GHR         | 7.23E-03     |
| GO_MF             | GO:0005096~GTPase activator activity           | 7           | ALDH1A1, AGAP11, RAP1GAP, TBC1D4, OPHN1, RGS16, STARID3 | 8.86E-03     |
| GO_MF             | GO:0008289~lipid binding                       | 10          | ALDH1A1, ALDH6A1, AR, MAP1LC3A, SDPR, OSBPL1A, FAAH, PPAR, CYTH3, PLEKHA2 | 1.06E-02     |
| GO_MF             | GO:0005534~phospholipid binding                | 6           | MAP1LC3A, SDPR, OSBPL1A, FAAH, CYTH3, PLEKHA2 | 1.44E-02     |
| GO_MF             | GO:0050662~coenzyme binding                    | 6           | ALDH6A1, ERO1LB, SORD, OGDHL, FOXRED2, WWOX | 1.57E-02     |
| Pathway           | hsa05200: Pathways in cancer                   | 8           | AR, BMP2, PGF, PPAR, FOXO1, ZBTB16, TCF7L1, DAPK1 | 1.28E-02     |

### B. Significant GO and KEGG terms enriched for the downregulated genes containing risk SNPs

| Category          | Term                                           | Gene number | Gene symbol                      | P-value      |
|-------------------|------------------------------------------------|-------------|----------------------------------|--------------|
| GO_BP             | GO:0022604~regulation of cell morphogenesis    | 20          | PALM, LST1, LZTS1, PTPRF, LIMK1, PLXNB2, CDH2, CDH4, TTL, TGFβ2, NRCAM, FYN, TIAM1,MYH14, CDC42EP3, ARAP1, RASA1, MYH10, CDC42EP5, FN1 | 4.41E-09     |
| GO_BP             | GO:0007155~cell adhesion                       | 48          | MPZL3, NRP2, CLDN4, POSTN, CD151, CDSN, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 2.56E-08     |
| Gene symbol | P-value |
|-------------|---------|
| MPZL3, NR2F2, CLDN9, POSTN, CD151, CD56, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 2.63E-08 |
| MPZL3, MYO6, ADAMTS14, LGAL3, SPOCK2, TNC, TGFB1, POSTN, CDH2, COL1A2, CDH1, ANXA2, COL1A1, ADAMTS2, DCBLD2, TNC, CD2, CDH3, TFNP51, TFNP54, TGFB2, COL22, RB2T, RON, CD4, MAP4, AFB1, AKB1, LYN, OLR1, CIB1, ILRN, ANXA1, CT1, MAF, BAX, CTBP1, PARP4, ROS1 | 1.63E-07 |
| MPZL3, NR2F2, CLDN9, POSTN, CD151, CD56, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 1.63E-07 |
| MPZL3, MYO6, ADAMTS14, LGAL3, SPOCK2, TNC, TGFB1, POSTN, CDH2, COL1A2, CDH1, ANXA2, COL1A1, ADAMTS2, DCBLD2, TNC, CD2, CDH3, TFNP51, TFNP54, TGFB2, COL22, RB2T, RON, CD4, MAP4, AFB1, AKB1, LYN, OLR1, CIB1, ILRN, ANXA1, CT1, MAF, BAX, CTBP1, PARP4, ROS1 | 1.63E-07 |
| MPZL3, NR2F2, CLDN9, POSTN, CD151, CD56, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 1.63E-07 |
| MPZL3, MYO6, ADAMTS14, LGAL3, SPOCK2, TNC, TGFB1, POSTN, CDH2, COL1A2, CDH1, ANXA2, COL1A1, ADAMTS2, DCBLD2, TNC, CD2, CDH3, TFNP51, TFNP54, TGFB2, COL22, RB2T, RON, CD4, MAP4, AFB1, AKB1, LYN, OLR1, CIB1, ILRN, ANXA1, CT1, MAF, BAX, CTBP1, PARP4, ROS1 | 1.63E-07 |
| MPZL3, NR2F2, CLDN9, POSTN, CD151, CD56, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 1.63E-07 |
| MPZL3, MYO6, ADAMTS14, LGAL3, SPOCK2, TNC, TGFB1, POSTN, CDH2, COL1A2, CDH1, ANXA2, COL1A1, ADAMTS2, DCBLD2, TNC, CD2, CDH3, TFNP51, TFNP54, TGFB2, COL22, RB2T, RON, CD4, MAP4, AFB1, AKB1, LYN, OLR1, CIB1, ILRN, ANXA1, CT1, MAF, BAX, CTBP1, PARP4, ROS1 | 1.63E-07 |
| MPZL3, NR2F2, CLDN9, POSTN, CD151, CD56, CTNNB1, NRCAM, PCDH1, WISP1, CD44 | 1.63E-07 |
| MPZL3, MYO6, ADAMTS14, LGAL3, SPOCK2, TNC, TGFB1, POSTN, CDH2, COL1A2, CDH1, ANXA2, COL1A1, ADAMTS2, DCBLD2, TNC, CD2, CDH3, TFNP51, TFNP54, TGFB2, COL22, RB2T, RON, CD4, MAP4, AFB1, AKB1, LYN, OLR1, CIB1, ILRN, ANXA1, CT1, MAF, BAX, CTBP1, PARP4, ROS1 | 1.63E-07 |
Table II. Continued.

B. Significant GO and KEGG terms enriched for the downregulated genes containing risk SNPs

| Category | Term | Gene number | Gene symbol | P-value |
|----------|------|-------------|-------------|---------|
| GO_CC    | GO:0044459–plasma membrane part | 119 | MICB, SRCIN1, PTGS2, ATP1B3, GABRB2, TLR2, CTNNB1, NRCAM, ST3GAL5, CD44, KCNK5 | 2.22E-11 |
| GO_CC    | GO:0005886–plasma membrane | 171 | MICB, ATP1B3, PTGS2, SRCIN1, GABRB2, TLR2, IQGAPI1, CTNNB1, NRCAM, MCOLN1 | 2.42E-10 |
| GO_CC    | GO:0031012–extracellular matrix | 34 | ADAMTS17, LTBP1, ADAMTS14, SPOCK2, TNC, ADAMTS5, POSTN, TGFβ2, TIMP1, LAMB3, CD44, FBN3, AGRN, FBN2, FGF1, LOXL1, FN1, LGALS3, COL13A1, PAPLN, SPARC, EMILIN2, COL5A2, ANXA2, ADAMTS9, BGN, NAV2, COL1A2, LAMC2, VCAN, COL1A1, DST, ADAMTS2, ADAMTS4 | 3.48E-09 |
| GO_CC    | GO:0005578–proteinaceous extracellular matrix | 32 | ADAMTS17, LTBP1, ADAMTS14, SPOCK2, TNC, ADAMTS5, POSTN, TIMP1, LAMB3, FBN3, AGRN, FBN2, FGF1, LOXL1, FN1, LGALS3, COL13A1, PAPLN, SPARC, EMILIN2, COL5A2, ANXA2, ADAMTS9, BGN, NAV2, COL1A2, LAMC2, VCAN, COL1A1, DST, ADAMTS2, ADAMTS4 | 7.79E-09 |
| GO_CC    | GO:0005887–integral to plasma membrane | 71 | MICB, LDLR, CLDN4, SLC20A1, ATP1B3, CORIN, GABRB2, TLR2, GABBR2, CD151, SDC4 | 2.00E-08 |
| Pathway  | hsa04514: Cell adhesion molecules (CAMs) | 19 | CLDN16, ICAM1, CLDN4, PTPRF, CDH2, HLA-B, CDH3, SDC4, CDH4, NRCAM, ITGA9, ITGB8, CD58, CLDN1, HLA-DPA1, VCAN, CNTNAP1, HLA-DOA, HLA-DRA | 1.16E-06 |
| Pathway  | hsa04512: ECM-receptor interaction | 15 | TNC, ITGA11, ITGA3, SDC4, COL5A2, ITGA9, LAMB3, CD44, ITGB8, ITGB6, COL1A2, LAMC2, AGRN, COL1A1, FN1 | 1.68E-06 |
| Pathway  | hsa04510: Focal adhesion | 22 | TNC, MET, ITGA11, ACTN1, ITGA3, COL5A2, FN1A, CTNNB1, ITGA9, LAMB3, CCND1, CCND3, CCND2, ITGB8, RASGRF1, FYN, ITGB6, COL1A2, LAMC2, COL1A1, SHC3, FN1 | 1.17E-05 |
| Pathway  | hsa05412: Arrhythmogenic right ventricular cardiomyopathy (ARVC) | 12 | ITGA9, DSG2, PKP2, ITGB8, ITGB6, ITGA11, CACNB1, ACTN1, ITGA3, CDH2, CACNA1C, CTNNB1 | 8.78E-05 |
| Pathway  | hsa04115: p53 signaling pathway | 11 | CCNE2, CCND1, TNFRSF10B, CCND3, CCND2, ZMAT3, BAX, SERPINE1, DDB2, FAS, PERP | 1.61E-04 |

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; SNP, single nucleotides polymorphism; BP, biological process; CC, cellular component; MF, molecular function.
A total of 7 PPI pairs containing the non-synonymous risk SNP loci in the interaction domains were identified. Particularly, the interaction domains involved in the interactions of FN1 and 5 other proteins (such as TNC) contained non-synonymous risk SNP loci. Prasad et al (34) hypothesized that an immunohistochemical panel containing FN1, Hector Battifora mesothelial cell 1 and galectin-3 may contribute to the diagnosis of thyroid tumors derived from follicular cells. Using reverse transcription-quantitative polymerase chain reaction, da Silveira Mitteldorf et al (35) demonstrated that FN1, MET, glutaminyl-peptide cyclotransferase, and UDP-galactose-4-epimerase were significantly upregulated in patients with PTC. TNC re-expression can be detected by immunohistochemistry in papillary and medullary thyroid carcinomas supplemented by the analysis of two TNC mRNA splice variants; thus TNC may be synthesized by tumor...
The present study was supported by the Qingdao Municipal Science and Technology Bureau (grant no. 13-1-4-166-jch).

Table III. Protein-protein interaction pairs containing the non-synonymous risk single nucleotide polymorphism loci in the interaction domains.

| Protein-1 | Protein-2 | Domain-1 | Domain-2 |
|-----------|-----------|----------|----------|
| TNC       | FN1       | PF00041  | PF00041  |
| FN1       | COL1A1    | PF00041  | PF01391  |
| FN1       | COL1A2    | PF00041  | PF01391  |
| COL1A1    | COL7A1    | PF01391  | PF01391  |
| FN1       | COL7A1    | PF00041  | PF01391  |
| FN1       | COL13A1   | PF00041  | PF01391  |
| MCM4      | MCM2      | PF00493  | PF00493  |

TNC, tenascin C; FN1, fibronectin 1; MCM, minichromosome maintenance; COL, collagen type.

Acknowledgements

The present study was supported by the Qingdao Municipal Science and Technology Bureau (grant no. 13-1-4-166-jch).

References

1. Carling T and Udelmans R: Thyroid cancer. Annu Rev Med 65: 125-137, 2014.
2. Vecchia CL, Malvezzi M, Bosetti C, Garavello W, Bertuccio P, Levi F and Negri E: Thyroid cancer mortality and incidence: A global overview. Int J Cancer 136: 2187-2195, 2015.
3. Dinets A, Hulchly M, Sohadiis A, Ghaderi M, Hooog A, Larsson C and Zedenius J: Clinical, genetic, and immunohistochemical characterization of adult thyroid carcinomas. J Clin Oncol 37: 93-99, 2011.
4. Costea V, Esposito R, Grubb J, Jeremiah SJ, Bethal JA, Demchick EP, Bogdanova TI, Tronko MD and Thomas GA: Investigation of loss of heterozygosity and SNP frequencies in the REL proto-oncogene are over-represented in patients with sporadic papillary thyroid carcinoma. J Med Genet 39: 260-265, 2002.
5. Stephens LA, Powell NG, Grubb J, Jeremiah SJ, Bethal JA, Demchick EP, Bogdanova TI, Tronko MD and Thomas GA: Investigation of loss of heterozygosity and SNP frequencies in the REL proto-oncogene are over-represented in patients with sporadic papillary thyroid carcinoma. J Med Genet 39: 260-265, 2002.
6. Salajegheh A, Smith RA, Kasem K, Gopalan V, Nassiri MR, William R and Lam AK: Single nucleotide polymorphisms and mRNA expression of VEGF-A in papillary thyroid carcinoma: Potential markers for aggressive phenotypes. Eur J Surg Oncol 37: 93-99, 2011.
7. Ciampi R, Knauf JA, Kerler R, Gandhi M, Zhu Z, Nikiforova MN, Rabes HM, Fagin JA and Nikiforov YE: Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. J Clin Invest 115: 94-101, 2005.
8. Wang X and Clark AG: Using next-generation RNA sequencing to identify imprinted genes. Heredity (Edinb) 113: 156-166, 2014.
9. de Klerk E, den Dunnen JT and ‘t Hoen PA: RNA sequencing: From tag-based profiling to resolving complete transcript structure. Cell Mol Life Sci 71: 3537-3551, 2014.
10. Costa V, Esposito R, Ziviello C, Sepe R, Bim LV, Cacciola NA, and Santoro M: Analysis of BRAF point mutation and RET/PTC rearrangement refines the fine-needle aspiration diagnosis of papillary thyroid carcinoma. J Clin Endocrinol Metab 95: 5175-5180, 2004.
11. Ciampi R, Knauf JA, Kerler R, Gandhi M, Zhu Z, Nikiforova MN, Rabes HM, Fagin JA and Nikiforov YE: Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. J Clin Invest 115: 94-101, 2005.
12. Wang X and Clark AG: Using next-generation RNA sequencing to identify imprinted genes. Heredity (Edinb) 113: 156-166, 2014.
18. Yang H and Wang K: Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10: 1556-1566, 2015.

19. Robinson MD, McCarthy DJ and Smyth GK: edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140, 2010.

20. Park BJ, Lord D and Hart J: Bias properties of Bayesian statistics in finite mixture of negative binomial regression models in crash data analysis. Accid Anal Prev 42: 741-749, 2010.

21. Haynes W: Benjamini-Hochberg Method. In: Encyclopedia of Systems Biology, Springer, pp78-78, 2013.

22. Ge H, Liu K, Juan T, Fang F, Newman M and Hoeck W: FusionMap: Detecting fusion genes from next-generation sequencing data at base-pair resolution. Bioinformatics 27: 1922-1928, 2011.

23. An J, Lai J, Saijanhar A, Batra J, Wang C and Nelson CC: J-Circos: An interactive Circos plotter. Bioinformatics 31: 1463-1465, 2015.

24. Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, Marygold S, Millburn G, Osumi-Sutherland D, Schroeder A, Seal K, et al.: FlyBase: Enhancing Drosophila gene ontology annotations. Nucleic Acids Res 37 (Database issue): D555-D559, 2009.

25. Koter A, Moriya Y, Tokimatsu T, Kanehisa M and Goto S: KEGG and GenoMNet, New Developments, Metagenomic Analysis. In: Encyclopedia of Metagenomics. Nelson KE (ed.) Springer, New York, NY, pp329-339, 2015.

26. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA: DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4: P3, 2003.

27. Goel R, Harsha HC, Pandey A and Prasad TS: Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. Mol Biosyst 8: 453-463, 2012.

28. Kohl M, Wisse S and Warscheid B: Cytoscape: Software for visualization and analysis of biological networks. Methods Mol Biol 696: 291-303, 2011.

29. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Quevri S, Sagardur-Vegar A, et al.: The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res 44: D279-D285, 2016.

30. Yellaboina S, Tasneem A, Zaykin DV, Raghavachari B and Johri R: DOMINE: A comprehensive collection of known and predicted domain-domain interactions. Nucleic Acids Res 39 (Database issue): D730-D735, 2011.

31. Ma S, Wang Q, Ma X, Wu L, Guo F, Ji H, Liu F, Zhao Y and Qin G: FoxP3 in papillary thyroid carcinoma induces NIS repression through activation of the TGF-β/Smad signaling pathway. Tumor Biol 120: 34-35, 2016.

32. Zhang J, Wang Y, Li D and Jing S: Notch and TGF-β/Smad3 pathways are involved in the interaction between cancer cells and cancer-associated fibroblasts in papillary thyroid carcinoma. Tumor Biol 35: 379-385, 2014.

33. Choe BK, Kim SK, Park HJ, Park HK, Kwon KH, Lim SH and Yim SV: Polymorphisms of TGFBR2 contribute to the progression of papillary thyroid carcinoma. Mol Cellular Toxicol 8: 1-8, 2012.

34. Prasad ML, Pellegrato NS, Huang Y, Nagaraja BH, de la Chapelle A and Kloos RT: Galectin-3, fibronectin-1, CITED-1, HBME1 and cytokeratin-19 immunohistochemistry is useful for the differential diagnosis of thyroid tumors. Mod Pathol 18: 48-57, 2005.

35. da Silveira Mittendorf CA, de Sousa-Canavez JM, Leite KR, Massicotte C and Camara-Lopes LH: FN1, GALE, MET and QPCT overexpression in papillary thyroid carcinoma: Molecular analysis using frozen tissue and routine fine-needle aspiration biopsy samples. Diagn Cytopathol 39: 556-561, 2011.

36. Tseleni-Balafouta S, Gakopoulou H, Fanourakis G, Voutsinas G, Balafoutas D and Patsouris E: Tenascin-C protein expression and mRNA splice variants in thyroid carcinoma. Exp Mol Pathol 80: 177-182, 2006.

37. Matsuse M, Takahashi M, Mitsutake N, Nishihara E, Hirokawa M, Kawaguchi T, Rogounovitch T, Saenko V, Bychkov A, Suzuki K, et al.: The FOXE1 and NKK2-1 loci are associated with susceptibility to papillary thyroid carcinoma in the Japanese population. J Med Genet 48: 645-648, 2011.

38. Ngan ES, Lang BH, Liu T, Shum CK, So MT, Lau DK, Leon TY, Cherny SS, Tsai SY, Lo CY, et al.: A germline mutation (A339V) in thyroid transcription factor-1 (TITF-1/NKX2.1) in patients with multinodular goiter and papillary thyroid carcinoma. J Natl Cancer Inst 101: 162-175, 2009.

39. Gilbert-Sirieix M, Makouki J, Kimura S, Talbot M, Caillou B, Massaad C and Massaad-Massade L: Wnt/β-catenin signaling pathway is a direct enhancer of thyroid transcription factor-1 in human papillary thyroid carcinoma cells. PLoS One 6: e22280, 2011.

40. Ambroziak M, Pachucki J, Stachlewksa-Nasfeter E, Nauman J and Nauman A: Disturbed expression of type 1 and type 2 iodothyronine deiodinase as well as titf1/nkx2.1 and Pax-8 transcription factor genes in papillary thyroid cancer. Thyroid 15: 1137-1146, 2005.

41. Kondo T, Nakazawa T, Ma D, Niu D, Momochiku K, Kawasaki T, Nakamura N, Yamane T, Kobayashi M and Katoh R: Epigenetic silencing of TFF-1/NKX2-1 through DNA hypermethylation and histone H3 modulation in thyroid carcinomas. Lab Invest 89: 791-799, 2009.