

## RESEARCH ARTICLE

# Integrated bioinformatics analysis and screening of hub genes in papillary thyroid carcinoma

Rong Fan<sup>1</sup>, Lijin Dong<sup>2</sup>, Ping Li<sup>3,4</sup>, Xiaoming Wang<sup>4\*</sup>, Xuewei Chen<sup>4\*</sup>

**1** Central Laboratory, Tianjin Xiqing Hospital, Tianjin, PR China, **2** Editorial Department of Education and Research Security Centre, Logistic University of Chinese People's Armed Police Force, Tianjin, PR China, **3** Southwest Medical University, Luzhou City, Sichuan Province, PR China, **4** Department of Operational Medicine, Tianjin Institute of Environmental and Operational Medicine, Tianjin, PR China

\* These authors contributed equally to this work.

\* [sisuo55123@sina.com](mailto:sisuo55123@sina.com) (XW); [chenxuewei11@sina.com](mailto:chenxuewei11@sina.com) (XC)

## Abstract

### Background

With the increasing incidence of papillary thyroid carcinoma (PTC), PTC continues to garner attention worldwide; however its pathogenesis remains to be elucidated. The purpose of this study was to explore key biomarkers and potential new therapeutic targets for, PTC.

### Methods

GEO2R and Venn online software were used for screening of differentially expressed genes. Hub genes were screened via STRING and Cytoscape, followed by Gene Ontology and KEGG enrichment analysis. Finally, survival analysis and expression validation were performed using the UALCAN online software and immunohistochemistry.

### Results

We identified 334 consistently differentially expressed genes (DEGs) comprising 136 upregulated and 198 downregulated genes. Gene Ontology enrichment analysis results suggested that the DEGs were mainly enriched in cancer-related pathways and functions. PPI network visualization was performed and 17 upregulated and 13 downregulated DEGs were selected. Finally, the expression verification and overall survival analysis conducted using the Gene Expression Profiling Interactive Analysis Tool (GEPIA) and UALCAN showed that LPAR5, TFPI, and ENTPD1 were associated with the development of PTC and the prognosis of PTC patients, and the expression of LPAR5, TFPI and ENTPD1 was verified using a tissue chip.

### Conclusions

In summary, the hub genes and pathways identified in the present study not only provide information for the development of new biomarkers for PTC but will also be useful for elucidation of the pathogenesis of PTC.

## OPEN ACCESS

**Citation:** Fan R, Dong L, Li P, Wang X, Chen X (2021) Integrated bioinformatics analysis and screening of hub genes in papillary thyroid carcinoma. PLoS ONE 16(6): e0251962. <https://doi.org/10.1371/journal.pone.0251962>

**Editor:** Giurato Giorgio, University of Salerno, ITALY

**Received:** July 27, 2020

**Accepted:** May 6, 2021

**Published:** June 11, 2021

**Copyright:** © 2021 Fan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The microarray data used to support the findings of this study were deposited in the Gene Expression Omnibus (GEO) repository (accession numbers: GSE3678, GSE33630, and GSE53157).

**Funding:** The study was supported by the National Natural Science Foundation of China (No.81201757).

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** PTC, Papillary thyroid cancer; GEO, Gene Expression Omnibus; DEGs, Differentially expressed genes; DAVID, Database for Annotation, Visualization, and Integrated Discovery; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Gene and Genome; BP, Biological process; CC, Cellular component; MF, Molecular function; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; MCODE, Molecular Complex Detection; GEPIA, Gene Expression Profiling Interactive Analysis; TCGA, The Cancer Genome Atlas; AKR1C3, Aldo-keto reductase family 1 member C3; LPAR5, Lysophosphatidic acid receptor 5; TFPI, Tissue factor pathway inhibitor; ENTPD1, Ectonucleoside triphosphate diphosphohydrolase 1; MAPK, Mitogen-activated protein kinase; BMP, Bone morphogenetic protein; TPA, 12-O-tetradecanoylphorbol-13-acetate; CLL, chronic lymphocytic leukemia; GC, gastric cancer; HNSCC, squamous cell carcinoma of the head and neck.

## Introduction

Over the past few years, thyroid diseases have garnered increasing attention and the incidence of thyroid cancer has markedly increased [1]. Thyroid cancer is divided into the following five types according to histogenesis and morphology: anaplastic, Hurthle cell, follicular, medullary, and papillary thyroid carcinoma, of which papillary thyroid carcinoma or papillary thyroid cancer (PTC) accounts for 80% of the overall incidence [2]. With the widespread use of neck ultrasound and fine-needle aspiration biopsy for diagnosis, the number of deaths attributable to thyroid cancer has significantly reduced [3]. However, the survival rate is affected by several factors, and the prognosis of PTC remains extremely poor despite the adoption of multiple treatment strategies such as thyroidectomy, radioiodine treatment, and chemotherapy [4]. Therefore, the early prevention and diagnosis of thyroid cancer remains a necessity and concern for doctors and scientists, and exploration of potential key biomarkers and novel therapeutic targets in PTC is imperative for doctors and patients alike.

With the large-scale application of high-throughput screening technology, we have identified several novel genes associated with disease initiation and progression and have achieved a comprehensive understanding of the molecular mechanisms of development of various tumors [5–8]. In the present study, we used bioinformatics methods to mine microarray data from the Gene Expression Omnibus (GEO) database, and analyzed the DEGs between PTC and normal tissues. Thereafter, we constructed a protein-protein interaction (PPI) network, performed functional enrichment and survival analysis, identified three hub genes, and discovered the biological processes and signaling pathways associated with PTC development. In conclusion, the integrated analysis provided insights for the development of new biomarkers for PTC, which may be valuable for conducting further research into the mechanisms of PTC as well as for use in clinical applications in diagnosis, prognosis, and therapy.

## Materials and methods

### Acquisition of microarray data

We obtained high-throughput gene expression profiles of PTC and normal thyroid tissues from the GEO database. The independent datasets GSE3678, GSE33630, and GSE53157 were selected, and all were based on the GPL570 platforms, including 7 PTC tissues and 7 normal tissues, 49 PTC tissues, 45 normal tissues, 7 PTC tissues, and 3 normal tissues.

### Screening of DEGs

The DEGs between PTC and normal tissues were screened using the GEO2R tool. GEO2R was used to compare and identify DEGs present in two or more sets of data in the GEO series. We used the adjusted *P*-value instead of the non-adjusted *P*-value to restrict the false-positive rate. The cut-off criteria of  $|\log_{2}FC| \geq 1$  and adjusted *P*-value  $< 0.05$  were considered statistically significant. The DEGs in the three datasets were screened using the Venn software online. DEGs with  $\log_{2}FC \geq 1$  were considered as upregulated genes, whereas DEGs with  $\log_{2}FC < -1$  were considered as downregulated genes.

### Enrichment analysis via GO and KEGG pathway

To characterize the functional roles of the DEGs, we used the Database for Annotation, Visualization, and Integrated Discovery (DAVID, version 6.8) [9] for GO enrichment analysis, which included biological process (BP), cellular component (CC), molecular function (MF), and KEGG pathway analysis with a cut-off *P*-value of  $< 0.05$ . Visualization of the enrichment

analysis was performed using the online platform Image GP (<http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html>).

### Construction of the PPI network and analysis of the module

The PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database to reveal the relationships of DEGs based on a minimum required interaction score = 0.4 [10]. The PPI network was illustrated and analyzed using the Cytoscape (version 3.6.1) software [11]. Additionally, the core modules of the PPI network were screened using MCODE (node score cutoff = 0.2, degree cutoff = 2, k-score = 2, and maximum depth = 100).

### Survival analysis and validation of hub gene expression

UALCAN was used to analyze the relationship between key gene expression and survival of patients with PTC, which is an established resource for analyzing transcriptome data of cancers based on The Cancer Genome Atlas (TCGA) [12]. The Gene Expression Profiling Interactive Analysis tool (GEPIA) [13] was used to analyze RNA expression data based on thousands of samples from TCGA and GTEx projects. Statistical significance was set at  $P < 0.05$ .

### Tissue samples and immunohistochemistry (IHC)

Human papillary thyroid cancer tissue microarray sections (HThyP120CS02) were obtained from Shanghai Outdo Biotech Co. Ltd. (Shanghai, China). The tissue samples were procured from 62 patients with papillary thyroid cancer, which consisted of 58 cancer tissues, 58 paracancerous tissues, 3 normal thyroid tissues, and 1 chronic lymphocytic thyroiditis tissue. The two-step EnVision method was used to perform immunohistochemical experiments, along with the use of different primary antibodies against LPAR5 (1:50). The study was conducted with the consent of human subjects and approved by the ethics committee of Tianjin Xiqing Hospital and the Shanghai Outdo Biotech Company (Shanghai, China). All experiments were conducted in accordance with the Declaration of Helsinki (1964). The slides were analyzed using the Image-Pro PLUS software program (Media Cybernetics, Inc. USA).

### Statistical analysis

Statistical analysis was performed using SPSS 22.0 and GraphPad Prism 8.0. Results are presented as mean  $\pm$  standard deviation. The Student's *t*-test was used to compare differences between two sample groups. Statistically significant changes are indicated with asterisks, where \*, \*\*, and \*\*\* represent  $P < 0.05$ ,  $P < 0.01$ , and  $P < 0.001$ , respectively.

## Results

### Identification of DEGs in papillary thyroid cancers

In this study, we selected and downloaded three GEO datasets and extracted DEG data based on the cut-off criteria. Data on a total of 63 papillary thyroid cancers and 55 normal thyroid tissues were obtained in this study. There were 636 DEGs in GSE3678, including 271 upregulated and 365 downregulated genes, 1241 DEGs in GSE33630, including 673 upregulated and 568 downregulated genes, and 884 DEGs in GSE53157, including 369 upregulated and 515 downregulated genes. Using the Venn diagram software, we detected a total of 334 common DEGs, which included 136 upregulated genes and 198 downregulated genes in PTC tissues (Table 1 and Fig 1).

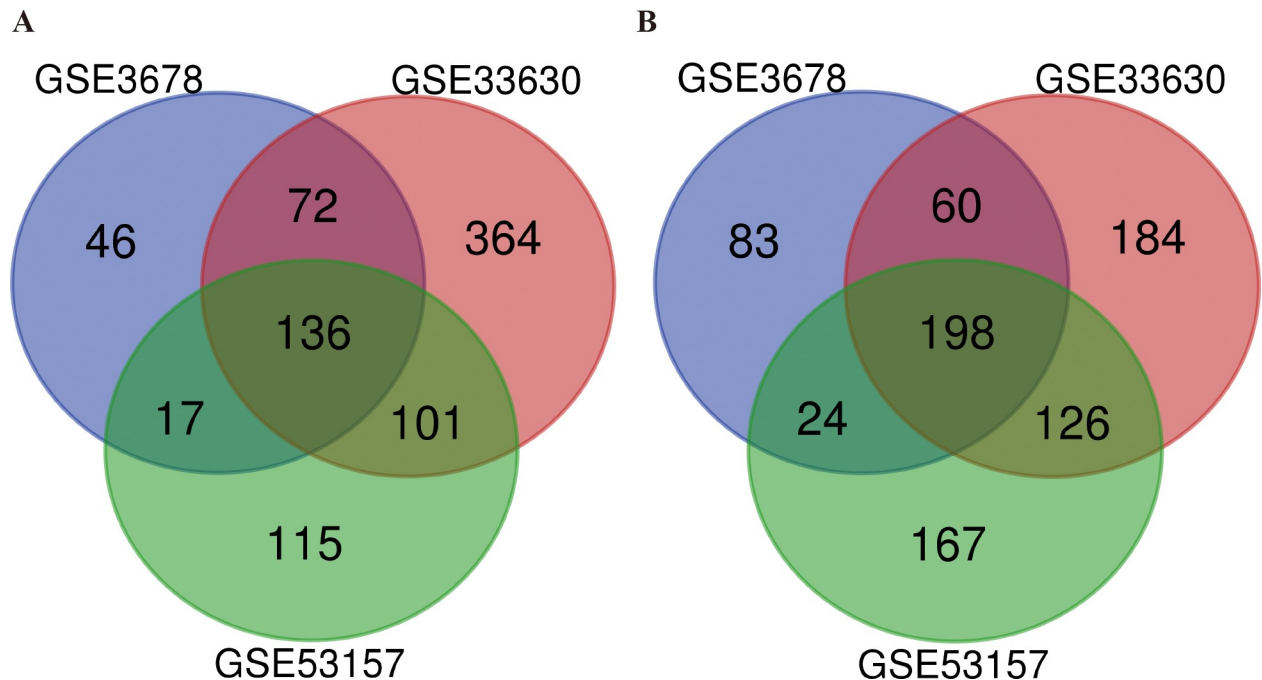
**Table 1. Screening of 334 commonly differentially expressed genes from three profile datasets, including 136 upregulated and 198 downregulated genes in PTC tissues, compared to normal thyroid tissues.**

DEGs	Gene symbol
Upregulated	CDH3, KLHDC8A, KCNJ2, GOLT1A, TNIK, PDZK1IP1, BID, RXRG, MPZL2, GDF15, FRMD3, DOCK9, MET, PTPRE, FAM20A, IGSF3, ALOX15B, ALOX5, EPS8, NGEF, ITGA3, DUSP6, PRSS2, CD55, NOD1, CFI, DDB2, DPP4, NPC2, DUSP4, CORO2A, ATP11A, ST8SIA4, LGALS3, PDLIM4, PLAG1, SYT12, RASD2, NFE2L3, TTC39B, DUSP5, LPAR5, TUSC3, FAXC, MYEF2, PDE5A, TACSTD2, PRR15, CDC42EP3, SNX25, RYR1, SPINT1, LRRK2, LOC101927705//P4HA2, NRP2, PRSS23, SIPA1L2, CHI3L1, CITED1, C19orf33, NR2F1-AS1, SERPINA1, TGFA, PLXNC1, LIPH, GGCT, LONRF2, TIAM1, FN1, SYTL1, RUNX2, SNX22, C4orf48, SYTL5, SERINC2, KRT19, ZCCHC12, LAMB3, LMO3, TENM1, SHROOM4, EMILIN2, ENTPD1, GALE, SDC4, METTL7B, HEY2, ETV5, ABCC3, NRCAM, LOC101928269//LOC100506403//RUNX1, GABRB2, SLC35F2, TNRC6C-AS1, TMPRSS4, SLC34A2, BICD1, STK32A, MYH10, RAB27A, TBC1D2, LRP4, CLDN1, LOC729461//FAM230B//FAM230C, TGFBR1, GABBR2, CCND1, NHSL2, MEGF9, CDH6, GALNT7, SREBF1, PTP4A3, TMC6, TPD52L1, PROS1, CAMK2N1, QPCT, SCEL, MCTP2, ETV1, KCNN4, KDELR3, PCSK2, COL8A1, ZMAT3, CYP1B1, HMGA2, LOC101928195//LOC100996643//MTHFD1L, LOC100507165, PSD3, AGR2, DTX4, CTSC, ARMCX3, XPR1
Downregulated	HBA2//HBA1, PKNOX2, MUM1L1, COL9A3, PBX4, IRS4, FGL2, MPPED2, DPT, TBC1D4, CITED2, ASXL3, MFAP4, FOXP2, FAM234B, DLG2, LAYN, TNS3, STXBPSL, SCARA5, TCEAL2, GPR83, RGS16, HGD, PAPSS2, GLT8D2, ID4, DGKI, EML1, GJB6, CDH16, ANO5, GHR, GNAI1, SCUBE3, LIFR, GRAMD2, MAGI2, CSGALNACT1, LYVE1, LOC101930400//AKR1C2, TMPRSS3, SOD3, GDF10, TFCP2L1, OTOS, EML6, STARD13, BCL2, KIAA1324, CXCL12, IPCEF1, RAPIGAP, PLEKHG4B, 2-Mar, AVPRIA, WDR72, SPX, IRS1, TSPAN7, PID1, SMAD9, BTBD11, C4orf47, ELMO1, NCAM1, DIO2, TPO, CASZ1, SORD, PQLC2L, SLC25A33, FCGBP, TBC1D8//RPL31, RYR2, PLA2R1, PKHD1L1, KCNIP4, GSTM3, ESRRG, CHCHD10, HBB, SERTM1, FHL1, RHOJ, SORBS2, MAFB, IP6K3, TFPI, C11orf74, RPS6KA6, MT1G, TMEM171, PKIA, PEG3, CCL21, ST7-AS1, SCN3A, CAPSL, PPARGC1A, DEPTOR, FBLN7, ERO1B, CRABP1, ZFPM2, RNF157-AS1, FHDC1, GNA14, MT1F, SHANK2, ANKS1B, BMP8A, COL23A1, TGFB3, ADH1B, AKR1C1, ANK2, NUAK2//AKIP1, KIZ, VLDLR, LMOD1, PRKCQ, DIO1, CCDC146, FGFR2, KIT, FABP4, ZDHHC11B//ZDHHC11, FAM167A, FAXDC2, ACACB, LOC440934, LOC646736, CYP7B1, RGS8, PGM5, MIR4683//FZD8, AADACP1, RNF150, AGPAT4, ZMAT4, TBX22, SLC16A2, TPPP, LOC101927137//KIAA1456, CTH, RASSF9, DDX25, TLE4, AKR1C3, SH3GL2, LOC101929480, DIRAS2, AGTR1, FMOD, ADAM22, LOC101060817//GCSH, DPY19L2, CLMN, SNCA, LRIG1, ANKRD18A, AIF1L, L3MBTL4, ADGRV1, RELN, CWH43, IGFBP1L, LOC100506558//MATN2, SLC26A7, WWOX, WASF3, UGT8, CUX2, SOX9-AS1, SLC4A4, TMEM178A, OGDHL, ADGRA3, DYNLRB2, TFF3, AOX1, GPM6A, CLCNKB, MDH1B, SAMD5, CFD, DPP6, TRIM58, INAFM2, FLRT1, MRO, TTC30A, EFEMP1, WFS1, SMOCC2, ANKRD37

<https://doi.org/10.1371/journal.pone.0251962.t001>

## GO and pathway enrichment analysis

DAVID was used for gene ontology functional annotation and biological pathway enrichment analysis of DEGs as a tool for the analysis of genes and proteins. We analyzed upregulated and downregulated genes using the DAVID software and identified 125 significant enrichment terms, including biological processes (BP, 79), molecular functions (MF, 22), and cellular components (CC, 24). As a common method adopted for defining genes and their corresponding RNA or protein products, GO analysis is usually performed to identify the unique biological characteristics of high-throughput transcriptomic or genomic data. BP, CC, and MF are the three main categories in the GO function annotation. In case of BP, DEGs were significantly enriched in cell adhesion, positive regulation of gene expression, BMP signaling pathway, transforming growth factor beta receptor signaling pathway, SMAD protein signal transduction, positive regulation of MAP kinase activity, ventricular septum morphogenesis, endothelial cell migration, and melanocyte differentiation. In case of MF, DEGs were significantly enriched in calcium ion binding, protease binding, glycoprotein binding, serine-type peptidase activity, RNA polymerase II transcription coactivator activity, growth factor binding, MAP kinase tyrosine/serine/threonine phosphatase activity, and semaphorin receptor activity. In case of CC, DEGs were remarkably enriched in the plasma membrane, extracellular space, cell



**Fig 1. Selection of 334 common DEGs from the three datasets (GSE3678, GSE33630, and GSE53157).** A. Venn diagram illustrating 136 upregulated DEGs ( $\log_{2}FC \geq 1$ ); B. Venn diagram illustrating 198 downregulated DEGs ( $\log_{2}FC < 1$ ). The number denotes the number of genes shared between/among subtypes or the number of unique genes.

<https://doi.org/10.1371/journal.pone.0251962.g001>

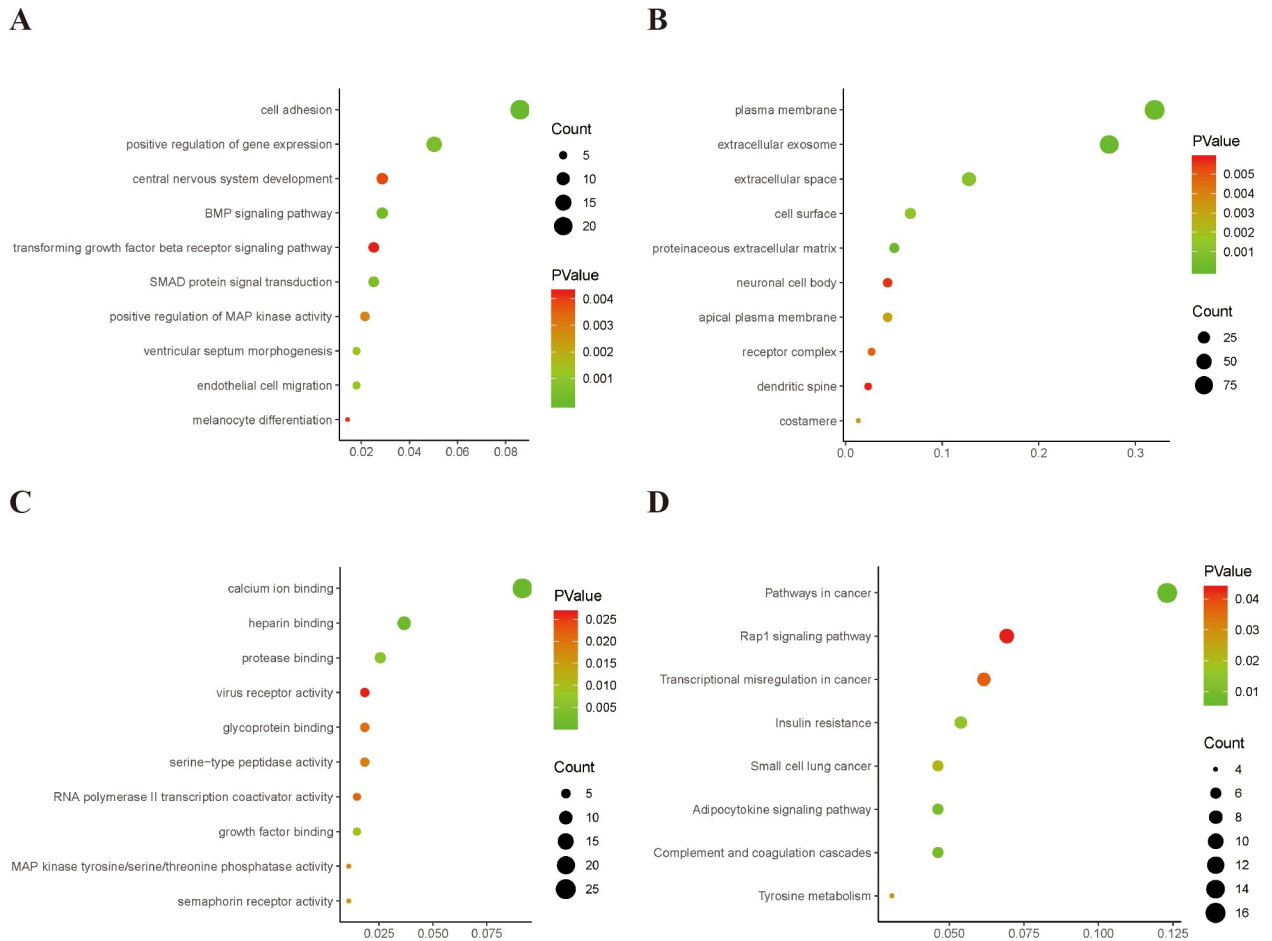
surface, proteinaceous extracellular matrix, neuronal cell body, apical plasma membrane, receptor complex, dendritic spine, and costamere. The top 10 GO terms are shown in Fig 2. Furthermore, KEGG is a collection of databases encompassing data of genomes, drugs, pathways, diseases, and chemical substances. KEGG pathway analysis showed that the DEGs were enriched in eight pathways, including pathways involved broadly in cancer, the Rap1 signaling pathway, transcriptional misregulation in cancer, insulin resistance, small cell lung cancer, adipocytokine signaling pathway, complement and coagulation cascades, and tyrosine metabolism (Fig 2).

### Analysis of the PPI network and module

Based on STRING and Cytoscape analysis, we constructed a PPI network complex with 241 nodes and 442 edges, including 101 upregulated and 137 downregulated genes. We then performed further analysis by applying Cytoscape MCODE plus and obtained 30 central nodes, including 17 upregulated genes and 13 downregulated genes (Figs 3 and 4, and Table 2).

### Selection of hub genes and validation of their expression level

We used GEPIA to identify the correlation between 30 central genes and the overall survival of patients with PTC. A total of 510 patients with PTC were included in the analysis. Patients with PTC were divided into high- and low-expression groups based on the median values of gene expression. The results suggested that the low expression levels of AKR1C3 ( $P = 0.00076$ ) and TFPI ( $P = 0.027$ ) were significantly correlated with shorter survival of patients with PTC, whereas the high expression levels of ENPD1 ( $P = 0.033$ ) and LPAR5 ( $P = 0.0049$ ) were remarkably related to the short survival of PTC patients (Fig 5). Furthermore, GEPIA was used to validate the expression levels of four hub genes in PTC and normal tissues. The results



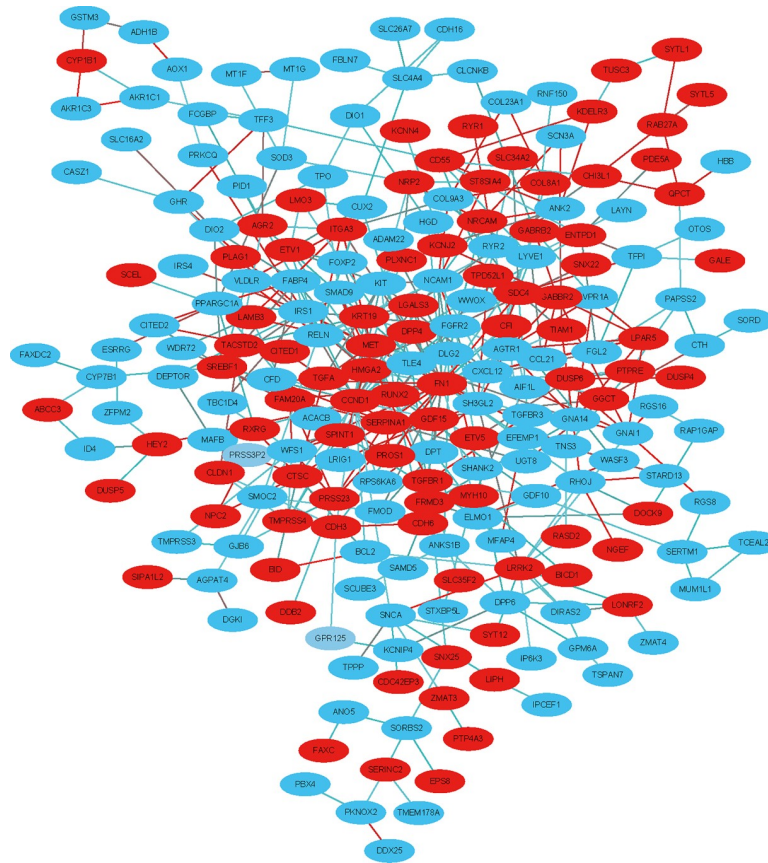
**Fig 2. The top 10 gene ontology (GO) and significantly enriched KEGG pathways.** A. BP; B. CC; C. MF; D. KEGG pathways. The X axis represents the enrichment levels; the Y axis represents the top 10 remarkably enriched categories. The color of the dot denotes the different *P*-value, and the size of the dot indicates the number of the candidate genes enriched in the GO and KEGG pathways.

<https://doi.org/10.1371/journal.pone.0251962.g002>

showed that the expression of all four genes was remarkably different in PTC samples compared to normal samples. ENTPD1 and LPAR5 expression levels were significantly upregulated, while the expression levels of AKR1C3 and TFPI were remarkably downregulated in PTC samples compared to normal samples (Fig 6). Moreover, UALCAN was used to analyze the expression levels of the four hub genes. We found that the expression levels of ENTPD1, LPAR5, and TFPI were remarkably different in PTC tissues compared to normal tissues, except for AKR1C3 (Fig 7). To verify the reliability of the prediction, we selected the LPAR5, TFPI and ENTPD1 protein and analyzed its expression using tissue microarray. The results indicated that LPAR5 expression was remarkably higher in papillary thyroid carcinoma tissues than that observed in the adjacent tissues (Fig 8), while there is no difference in the expression of TFPI and ENTPD1 between PTC tissues and the adjacent tissues (S1 Fig).

### Discussion

In this study, we analyzed gene expression profiles obtained from three GEO datasets (GSE 3678, GSE 33630, and GSE 53157 by GEO2R and identified 136 upregulated and 198 downregulated genes, for a total of 334 DEGs. Particularly, the data sets selected herein were obtained from the same data platform, and the purpose was to ensure data uniformity and reliability. 63

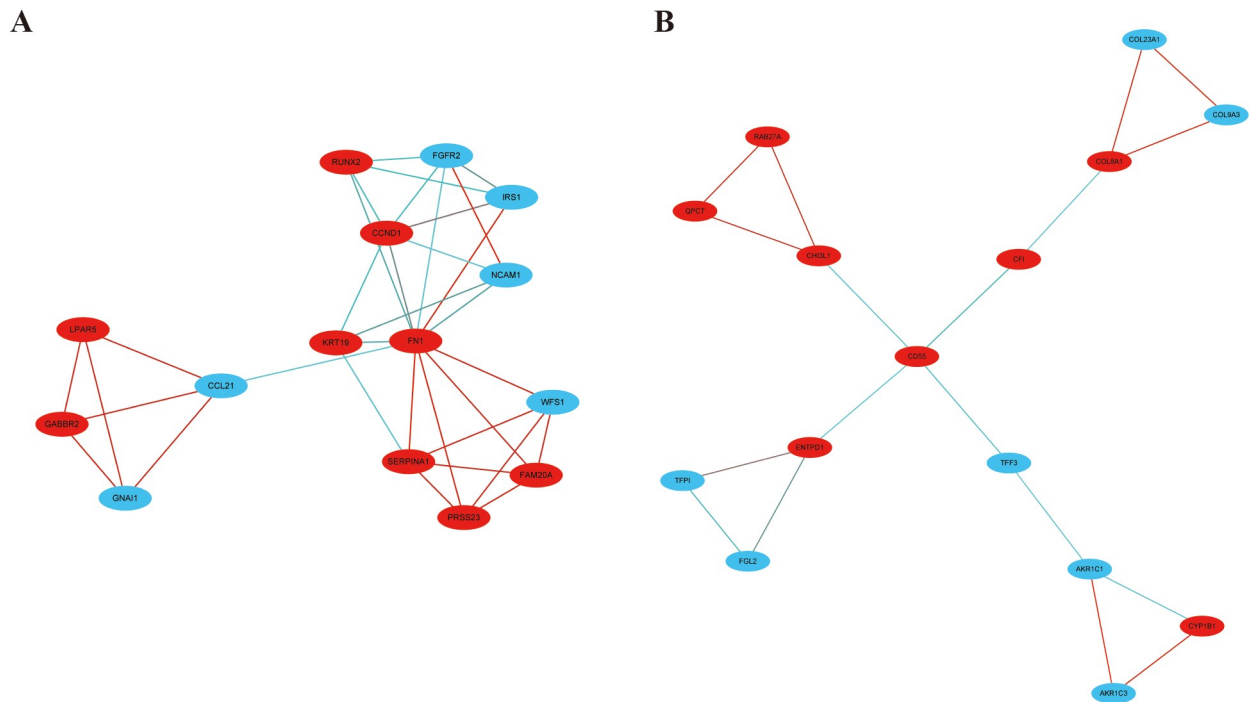


**Fig 3. PPI network construction using STRING and Cytoscape.** Each node indicates a protein module; the edges represent protein interactions; red circles denote upregulated DEGs and blue circles represent downregulated DEGs.

<https://doi.org/10.1371/journal.pone.0251962.g003>

papillary thyroid cancer specimens and 55 normal thyroid specimens were enrolled in this study. Gene Ontology and KEGG pathway enrichment analysis suggested that these DEGs were remarkably enriched in certain pathways, including pathways in cancer, TGFβ receptor signaling pathway, growth factor binding, and were markedly associated with BMP, SMAD, and MAPK signal activity. Thereafter, we constructed a PPI network using the STRING and Cytoscape software and performed screening for 30 vital hub genes from the PPI network complex via MCODE module analysis. Furthermore, we found that 4 of the 30 genes were markedly associated with survival via GEPIA analysis. Expression validation based on GEPIA and UALCAN showed that there was a significant difference in the expression of ENTPD1, LPAR5, and TFPI in PTC and normal tissues, suggesting that these genes might play critical roles in the tumorigenesis and development of PTC.

Previous research has indicated that the three genes identified in our study are involved in the initiation and development of tumors. Lysophosphatidic acid (LPA) is a biologically active mediator that affects cellular functions, such as regulation of cell proliferation, transcellular migration, differentiation, morphogenesis, and prevention of apoptosis [14, 15]. LPAR5 is a member of the G protein-coupled transmembrane receptor that establishes interaction with LPA [16, 17]. The cells subjected to treatment with 12-O-tetradecanoylphorbol-13acetate (TPA) showed higher motility than control cells, whereas LPAR5 knockdown reversed this phenomenon [18]. Recent studies on PTC have shown that LPAR5 is associated with progression and overall survival in thyroid cancer, which is consistent the findings presented in our



**Fig 4. Module analysis by MCODE using Cytoscape.** A. module1; B. module2. Red circles represent upregulated DEGs and blue circles denote downregulated DEGs.

<https://doi.org/10.1371/journal.pone.0251962.g004>

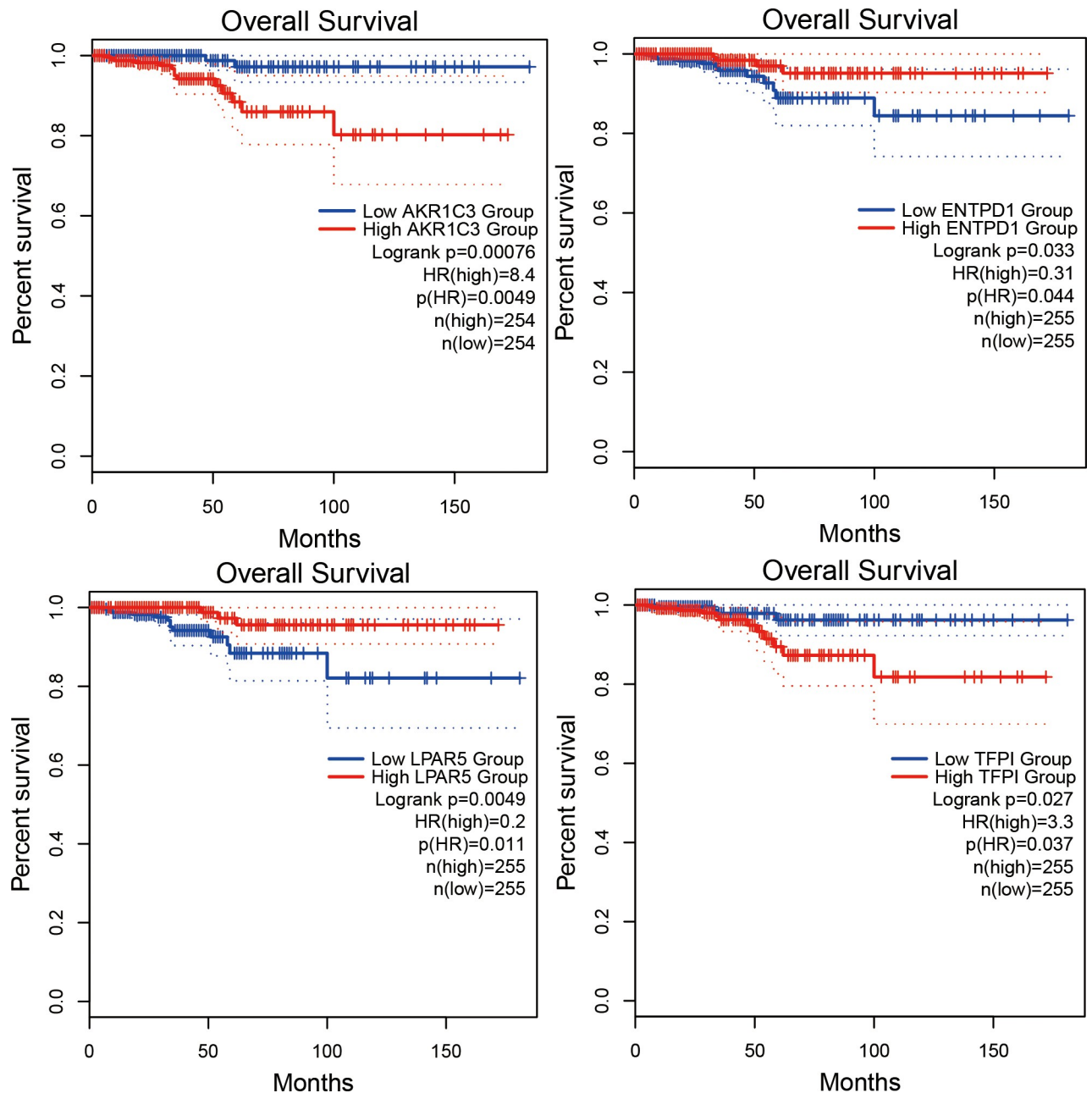
study [19–21]. Tissue factor pathway inhibitor (TFPI) is an endogenous inhibitor of tissue factor-induced blood coagulation, and its expression has been demonstrated in smooth muscle cells, monocytes, platelets, and several breast cancer cell lines [22–24]. Wang et al. suggested that downregulation of TFPI might result in tumor cell growth and migration, and the suppression of TFPI by hypoxia microenvironment might be one of the supervisory mechanisms by which hypoxia could promote angiogenesis and tumor growth [25]. miR-500 inhibition could suppress the proliferation and invasion of prostate cancer cells and tumorigenicity in vivo, while TFPI knockdown reversed these effects [26]. The expression level of TFPI in luminal-A breast cancer patients was significantly lower than that in healthy volunteers [27]. Zar-ychta et al. claimed that TF seemed to be a tumor-promoting factor, while TFPI exhibited tumor suppressor properties [28]. Erem et al. found a significant decrease in TFPI levels in patients with hyperthyroidism [29]. ENTPD1, which encodes the protein CD39, is extremely important for the production of immunosuppressive adenosine [30]. Mosaad Zaki et al. confirmed that the expression level of ENTPD1 on CD4+ T helper cells in chronic lymphocytic leukemia (CLL) patients was significantly higher than that in the controls, and ENTPD1 and CD4 expression levels were remarkably expressed in high-risk CLL patients [31]. Cai et al. found that ENTPD1 overexpression was a predictor of poor prognosis in GC patients after subsection to radical gastric cancer (GC) resection [32]. Pathological research has shown that

**Table 2. Selection of 30 central genes from PPI network, including 17 upregulated and 13 downregulated genes, by using the STRING and Cytoscape software.**

DEGs	Gene symbol
<b>Upregulated</b>	CCND1, FAM20A, FN1, PRSS23, LPAR5, SERPINA1, KRT19, GABBR2, RUNX2, CHI3L1, CFI, QPCT, CD55, CYP1B1, RAB27A, ENTPD1, COL8A1
<b>Downregulated</b>	GNAI1, NCAM1, CCL21, IRS1, WFS1, FGFR2, FGL2, AKR1C3, AKR1C1, TFF3, COL23A1, TFPI, COL9A3

<https://doi.org/10.1371/journal.pone.0251962.t002>

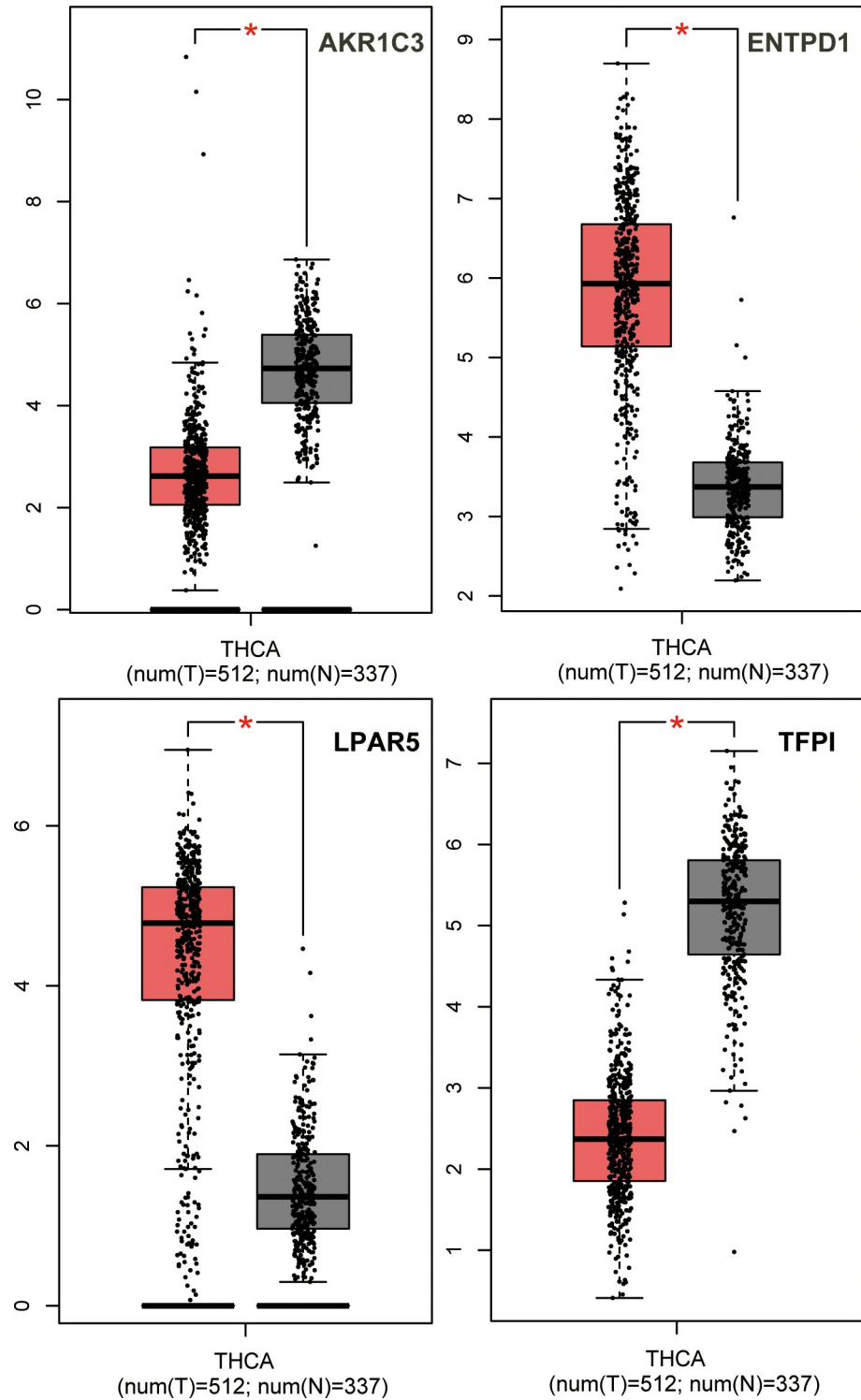




**Fig 5. The association between the expression of the four hub genes and the overall survival of patients with PTC via GEPIA analysis.** Red curve denotes high expression level, while blue curve denotes low expression level.

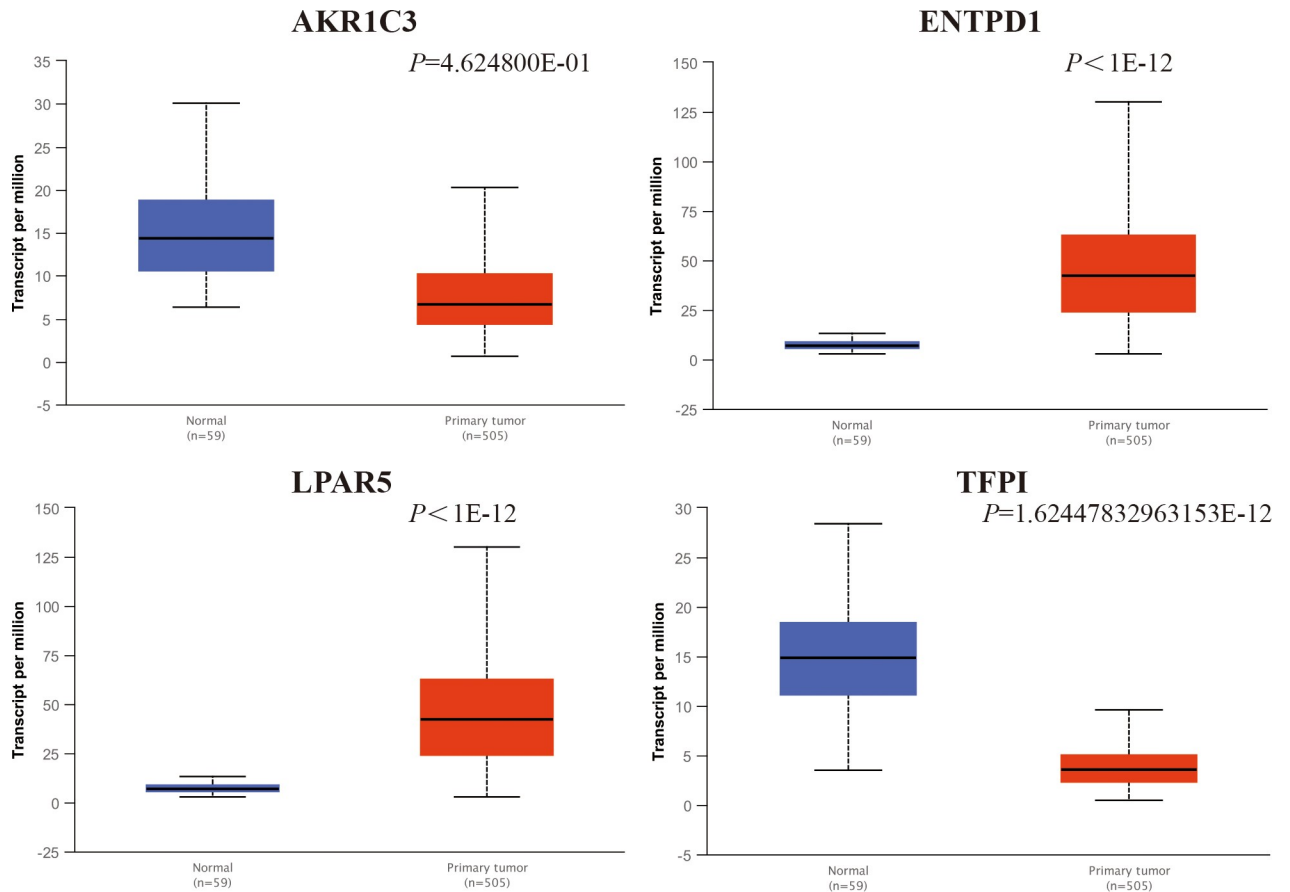
<https://doi.org/10.1371/journal.pone.0251962.g005>

the expression of ENTPD1 in head and neck squamous cell carcinoma (HNSCC) is positively correlated with tumor stage, which may predict a poor prognosis [33]. Interestingly, Sun et al. found that ENTPD1 deletion promoted the development of both induced and spontaneous autochthonous liver cancer in mice [34]. These results suggest that the development of immunotherapy targeting CD39 may be promising. Thus far, there have been a few reports on the relationship between ENTPD1 and thyroid cancer. Thus, the specific functions of ENTPD1 in thyroid cancer should be further explored. However, further experiments are warranted to explore the pathogenesis and molecular mechanisms of the hub genes in PTC.



**Fig 6. Analysis of hub gene expression levels with GEPIA analysis.** The transcript expression comparison of four genes in normal (n = 59) and tumor (n = 505) tissues via GEPIA.

<https://doi.org/10.1371/journal.pone.0251962.g006>

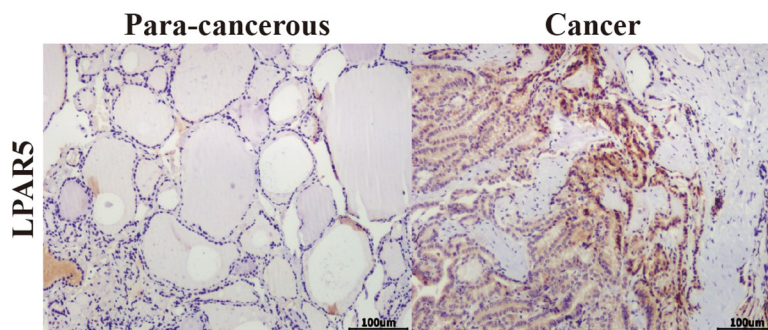


**Fig 7. Analysis of hub gene expression levels with UALCAN analysis.** The transcript expression comparison of four genes in normal (n = 59) and tumor (n = 505) tissues via UALCAN.

<https://doi.org/10.1371/journal.pone.0251962.g007>

## Conclusions

In summary, our integrated bioinformatics study presented three hub genes LPAR5, and pathways associated with PTC which might be a reliable as potential biomarkers, and might provide new insights into the diagnosis, prognosis, and target therapy for PTC.



**Fig 8. Representative IHC for LPAR5 expression in tissue microarrays (scale bar, 100 µm).** The categories of the expression of LPAR5 immunostaining in the papillary thyroid carcinoma at 100× magnification. Left: para-cancerous tissue; right: cross-section of a tumor obtained from PTC patient shows intense cytoplasmic staining.

<https://doi.org/10.1371/journal.pone.0251962.g008>

## Supporting information

**S1 Fig.**  
(TIF)

## Acknowledgments

The visualization of GO analysis, including BP, MF, and CC of DEGs, was performed using the online software ImageGP (<http://www.ehbio.com/ImageGP/index.php/Home/Index/index>).

## Author Contributions

**Conceptualization:** Rong Fan, Xiaoming Wang.

**Data curation:** Lijin Dong.

**Methodology:** Ping Li.

**Writing – original draft:** Xiaoming Wang.

**Writing – review & editing:** Xuewei Chen.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA: a cancer journal for clinicians*. 2017; 67(1):7–30.
2. Tuttle RM, Ball DW, Byrd D, Dilawari RA, Doherty GM, Duh QY, et al. Thyroid carcinoma. *Journal of the National Comprehensive Cancer Network: JNCCN*. 2010; 8(11):1228–74. <https://doi.org/10.6004/jnccn.2010.0093> PMID: 21081783
3. La Vecchia C, Malvezzi M, Bosetti C, Garavello W, Bertuccio P, Levi F, et al. Thyroid cancer mortality and incidence: a global overview. *International journal of cancer*. 2015; 136(9):2187–95. <https://doi.org/10.1002/ijc.29251> PMID: 25284703
4. Schlumberger M, Sherman SI. Approach to the patient with advanced differentiated thyroid cancer. *European journal of endocrinology*. 2012; 166(1):5–11. <https://doi.org/10.1530/EJE-11-0631> PMID: 21890651
5. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science (New York, NY)*. 2013; 339(6127):1546–58.
6. Zhao P, Hu W, Wang H, Yu S, Li C, Bai J, et al. Identification of differentially expressed genes in pituitary adenomas by integrating analysis of microarray data. *International journal of endocrinology*. 2015; 2015:164087. <https://doi.org/10.1155/2015/164087> PMID: 25642247
7. Gao X, Wang X, Zhang S. Bioinformatics identification of crucial genes and pathways associated with hepatocellular carcinoma. *Bioscience reports*. 2018; 38(6). <https://doi.org/10.1042/BSR20181441> PMID: 30341252
8. Sun C, Cheng X, Wang C. Gene expression profiles analysis identifies a novel two-gene signature to predict overall survival in diffuse large B-cell lymphoma. 2019; 39(1).
9. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009; 4(1):44–57. <https://doi.org/10.1038/nprot.2008.211> PMID: 19131956
10. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. 2017; 45(D1): D362–d8.
11. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003; 13(11):2498–504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
12. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia (New York, NY)*. 2017; 19(8):649–58. <https://doi.org/10.1016/j.neo.2017.05.002> PMID: 28732212

13. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research*. 2017; 45(W1):W98–w102. <https://doi.org/10.1093/nar/gkx247> PMID: 28407145
14. Furui T, LaPushin R, Mao M, Khan H, Watt SR, Watt MA, et al. Overexpression of edg-2/vzg-1 induces apoptosis and anoikis in ovarian cancer cells in a lysophosphatidic acid-independent manner. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 1999; 5(12):4308–18.
15. Contos JJ, Ishii I, Fukushima N, Kingsbury MA, Ye X, Kawamura S, et al. Characterization of lpa(2) (Edg4) and lpa(1)/lpa(2) (Edg2/Edg4) lysophosphatidic acid receptor knockout mice: signaling deficits without obvious phenotypic abnormality attributable to lpa(2). *Molecular and cellular biology*. 2002; 22(19):6921–9. <https://doi.org/10.1128/MCB.22.19.6921-6929.2002> PMID: 12215548
16. An S, Bleu T, Hallmark OG, Goetzl EJ. Characterization of a novel subtype of human G protein-coupled receptor for lysophosphatidic acid. *The Journal of biological chemistry*. 1998; 273(14):7906–10. <https://doi.org/10.1074/jbc.273.14.7906> PMID: 9525886
17. Lee CW, Rivera R, Gardell S, Dubin AE, Chun J. GPR92 as a new G12/13- and Gq-coupled lysophosphatidic acid receptor that increases cAMP, LPA5. *The Journal of biological chemistry*. 2006; 281(33):23589–97. <https://doi.org/10.1074/jbc.M603670200> PMID: 16774927
18. Fukushima K, Takahashi K, Kurokawa A, Ishimoto K, Otagaki S, Minami K, et al. Involvement of LPA receptor-5 in the enhancement of cell motile activity by phorbol ester and anticancer drug treatments in melanoma A375cells. *Biochemical and biophysical research communications*. 2018; 496(1):225–30. <https://doi.org/10.1016/j.bbrc.2018.01.030> PMID: 29309788
19. Tang J, Kong D, Cui Q, Wang K, Zhang D, Yuan Q, et al. Bioinformatic analysis and identification of potential prognostic microRNAs and mRNAs in thyroid cancer. *PeerJ*. 2018; 6:e4674. <https://doi.org/10.7717/peerj.4674> PMID: 29740512
20. Liu L, He C, Zhou Q, Wang G, Lv Z, Liu J. Identification of key genes and pathways of thyroid cancer by integrated bioinformatics analysis. 2019; 234(12):23647–57.
21. Zhang S, Wang Q, Han Q, Han H, Lu P. Identification and analysis of genes associated with papillary thyroid carcinoma by bioinformatics methods. *Bioscience reports*. 2019; 39(4). <https://doi.org/10.1042/BSR20190083> PMID: 30872410
22. Stavik B, Skretting G, Aasheim HC, Tinholt M, Zernichow L, Sletten M, et al. Downregulation of TFPI in breast cancer cells induces tyrosine phosphorylation signaling and increases metastatic growth by stimulating cell motility. *BMC cancer*. 2011; 11:357. <https://doi.org/10.1186/1471-2407-11-357> PMID: 21849050
23. Stavik B, Tinholt M, Sletten M, Skretting G, Sandset PM, Iversen N. TFPIalpha and TFPIbeta are expressed at the surface of breast cancer cells and inhibit TF-FVIIa activity. *Journal of hematology & oncology*. 2013; 6:5.
24. Ravindranath TM, Goto M, Iqbal O, Florian-Kujawski M, Hoppensteadt D, Hammadeh R, et al. Plasma thrombin activatable fibrinolysis inhibitor and tissue factor pathway inhibitor changes following sepsis. *Clinical and applied thrombosis/hemostasis: official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis*. 2007; 13(4):362–8. <https://doi.org/10.1177/1076029607305580> PMID: 17911187
25. Wang T, Gilkes DM, Takano N, Xiang L, Luo W, Bishop CJ, et al. Hypoxia-inducible factors and RAB22A mediate formation of microvesicles that stimulate breast cancer invasion and metastasis. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(31):E3234–42. <https://doi.org/10.1073/pnas.1410041111> PMID: 24938788
26. Cai B, Chen W, Pan Y, Chen H, Zhang Y, Weng Z, et al. Inhibition of microRNA-500 has anti-cancer effect through its conditional downstream target of TFPI in human prostate cancer. 2017; 77(10):1057–65. <https://doi.org/10.1002/pros.23361> PMID: 28631332
27. Tinholt M, Vollan HK, Sahlberg KK, Jernstrom S, Kaveh F, Lingjaerde OC, et al. Tumor expression, plasma levels and genetic polymorphisms of the coagulation inhibitor TFPI are associated with clinicopathological parameters and survival in breast cancer, in contrast to the coagulation initiator TF. *Breast cancer research: BCR*. 2015; 17:44. <https://doi.org/10.1186/s13058-015-0548-5> PMID: 25882602
28. Zarychta E, Rhone P, Bielawski K, Rosc D, Szot K, Zdunska M, et al. Elevated plasma levels of tissue factor as a valuable diagnostic biomarker with relevant efficacy for prediction of breast cancer morbidity. *Journal of physiology and pharmacology: an official journal of the Polish Physiological Society*. 2018; 69(6). <https://doi.org/10.26402/jpp.2018.6.06> PMID: 30802214
29. Erem C, Ucuncu O, Yilmaz M, Kocak M, Nuhoglu I, Ersoz HO. Increased thrombin-activatable fibrinolysis inhibitor and decreased tissue factor pathway inhibitor in patients with hypothyroidism. *Endocrine*. 2009; 35(1):75–80. <https://doi.org/10.1007/s12020-008-9116-4> PMID: 18958631

30. Allard B, Longhi MS, Robson SC, Stagg J. The ectonucleotidases CD39 and CD73: Novel checkpoint inhibitor targets. *Immunological reviews*. 2017; 276(1):121–44. <https://doi.org/10.1111/imr.12528> PMID: 28258700
31. Mosaad Zaki E, Mohamed Zahran A, Abdelazeem Metwaly A, Hafez R, Hussein S, Elaiw Mohammed A. Impact of CD39 expression on CD4+ T lymphocytes and 6q deletion on outcome of patients with chronic lymphocytic leukemia. *Hematology/oncology and stem cell therapy*. 2019; 12(1):26–31. <https://doi.org/10.1016/j.hemonc.2018.09.002> PMID: 30336122
32. Cai XY, Wang XF, Li J, Dong JN, Liu JQ, Li NP, et al. High expression of CD39 in gastric cancer reduces patient outcome following radical resection. *Oncology letters*. 2016; 12(5):4080–6. <https://doi.org/10.3892/ol.2016.5189> PMID: 27895775
33. Mandapathil M, Boduc M, Roessler M, Guldner C, Walliczek-Dworschak U, Mandic R. Ectonucleotidase CD39 expression in regional metastases in head and neck cancer. *Acta oto-laryngologica*. 2018; 138(4):428–32. <https://doi.org/10.1080/00016489.2017.1405278> PMID: 29172836
34. Sun X, Han L, Seth P, Bian S, Li L, Csizmadia E, et al. Disordered purinergic signaling and abnormal cellular metabolism are associated with development of liver cancer in Cd39/ENTPD1 null mice. *Hepatology (Baltimore, Md)*. 2013; 57(1):205–16. <https://doi.org/10.1002/hep.25989> PMID: 22859060