FBLN5 Is an Underlying Common Tumor Suppressor in Breast Cancer and Thyroid Cancer

Zhao-min XIE

the Cancer Hospital of Shantou University Medical College

Ying-sheng XIAO

Shantou Central Hospital

Chun-yan XU

the Cancer Hospital of Shantou University Medical College

Qin XIE

the Cancer Hospital of Shantou University Medical College

Wen-de WANG

the Cancer Hospital of Shantou University Medical College

Zheng-hao WANG

The First Affiliated Hospital of Shantou University Medical College

Yuan-ke LIANG

The First Affiliated Hospital of Shantou University Medical College

Hao-yu LIN

The First Affiliated Hospital of Shantou University Medical College

De ZENG (✉ dezeng@stu.edu.cn)

the Cancer Hospital of Shantou University Medical College

Research Article

Keywords: breast cancer, thyroid cancer, WGCNA, FBLN5, PI3K/AKT, second primary cancer

Posted Date: January 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1162202/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Breast cancer (BC) patients have a greater risk of developing thyroid cancer (TC) than the general population. Similarly, TC patients are more likely to develop BC, suggesting an underlying common etiology. In this study, we sought to identify the potential cross-talking pathway and related molecular mechanisms conferring to the sequential development of BC and TC.

Methods: We first used Multiple Primary-Standardized Incidence Ratios (MP-SIR) Program of SEER*Stat to calculate SIR to confirm the relationship between BC and TC. Then the RNA-seq was downloaded from The Cancer Genome Atlas (TCGA). And we built a co-expression network via Weighted Gene Co-expression Network Analysis (WGCNA) and obtained the most significant modules. The key genes were obtained by differential gene expression (DGE) analysis and WGCNA analysis. Furthermore, String database and Cytoscape software were used to construct protein-protein interactions (PPI), and defined the maximum Maximal Clique Centrality (MCC) value as hub gene. Then we performed prognosis analysis on the hub genes and obtained the prognostic genes of BC and TC. Finally, gene set enrichment analysis (GSEA) was used to investigate the molecular pathways associated with prognostic gene expressed both in BC and TC.

Results: From the SEER database, we found that the risk of developing BC in TC patients was SIR 1.12, 95% CI [1.07, 1.18], and the risk of developing TC in BC patients was SIR 1.29, 95% CI [1.23, 1.36]. Fifty-nine key genes obtained by differential expression analysis and WGCNA identify that PI3K/AKT was the most enriched pathway in BC and TC. In addition, the Recombinant Fibulin 5 (FBLN5) was shown to be of significant prognostic value for both BC and TC and was down-regulated in BC and TC tissues. GSEA demonstrated that FBLN5 enrichment pathways associated with BC and TC mainly included: B cell receptor signaling pathway, steroid hormone biosynthesis, and pathways in cancer.

Conclusions: The PI3K/AKT signaling is most co-enriched pathway in BC and TC. FBLN5 is the most relevant prognostic gene and an underlying common tumor suppressor in both BC and TC, with downstream pathways involving immunity, hormone biosynthesis and carcinogenesis.

1 Introduction

According to newly released data, BC and TC were among the most prevalent malignancies in women. There were about 2.1 million new cases of BC in women worldwide, accounting for almost a quarter of all cancer cases in women in 2018. In addition, around 567,000 cases were diagnosed with TC and ranked ninth in cancer incidence worldwide, with a global female incidence rate of 10.2 per 100,000 people, three times that of male[1]. A number of studies have reported that the risk of TC increases after BC, similarly, the risk of BC increases after TC[2-5]. Therefore, BC and TC might share certain etiologies in common, and studies have shown that obesity, genetic susceptibility, hormone factors and certain environmental pollutants were all risk factors for BC and TC[6]. However, their underlying common mechanisms of action remain unknown. A few epidemiological studies have suggested a link between BC and TC, and we

used the SEER database to verify that BC and TC have a link, but until recently, no study have proven it at the molecular level. Therefore, we aim to investigate whether there is a genetic link between BC and the TC.

SEER database is an authoritative neoplasm statistics database of the United States. It records information on the incidence, mortality and morbidity of millions of malignancies in several counties in the United States. The cancer information in the database is unified and standardized by SEER*Stat software. We used the SEER*Stat software program and selected patients between 1973 and 2015 to obtain the SIR of TC in BC patients and SIR of BC in TC patients, and to verify the relationship between BC and TC.

With the development of bioinformatics and genomics technology, gene expression profiling can be used to explore the molecular mechanism of disease etiology[7]. DGE analysis of transcriptomics provided an effective method to study the molecular mechanisms of genome regulation[8]. WGCNA is a kind of bioinformatics analysis method for constructing gene co-expression networks, aiming at obtaining co-expression gene modules. It can cluster genes with similar expression patterns, explore associations between gene networks and phenotypes of interest, build a bridge between sample characteristics and changes in gene expression, and ultimately identify key genes in the disease pathway[9][10, 11]. The joint of WGCNA and DGE analysis can obtain highly correlated genes. Therefore, In this study, we combined DGE with WGCNA analysis to obtain highly correlated genes.

In this paper, we used the SEER database to determine the incidence of secondary malignancies with BC as TC or TC as BC. Then, the RNA-seq data of BC and TC were obtained from TCGA data set, and the key genes were obtained by DGE analysis and WGCNA analysis. Functional enrichment analysis was used to explore the commonality of BC and TC. We constructed PPI to select the hub genes. Subsequently, hub genes related to the prognosis of BC and TC were screened. Finally, GSEA of prognostic genes was performed to explore the molecular pathways related to the expression of prognostic genes in BC and TC. This study might provide a new perspective to a possible shared molecular mechanism of the development of BC and TC.

2 Materials And Methods

The design and process of this article were shown in Figure 1.

2.1 Data collection, processing and differentially expressed genes (DEGs) screening

The RNA-Seq data of BC and TC were obtained from TCGA database, which is the maximum cancer gene information database at present and the most abundant and authoritative database in cancer research. Bioconductor package in R includes a limma package that can be used to analyze biochip technology. Limma can read the raw data, control the quality of the data, preprocess the data and analyze the

differences[12]. We used limma package to acquire DEGs of BC and TC. The cut-off standard were $|\log FC| \geq 1.0$ and $\text{adj. } P < 0.05$ (adjusted p value), and the DEGs of BC and TC were visualized as volcano plot adopting ggplot2 package in R.

2.2 Data extraction from the SEER database

The MP-SIR Session of SEER* Stat software can be used to calculate the risk of developing a second primary cancer in people who have already been diagnosed with cancer, compared with the general population. We obtained the information of BC and TC from the SEER database. Patients from 1973 to 2015 were selected, and then the MP-SIR Session of SEER* Stat software was performed to obtain the SIR of TC patients with the second cancer and SIR of BC patients with the second cancer. The SIR is equal to the observed count divided by the expected count, and the 95% confidence interval (CI) can be obtained by calculating the assumed Poisson distribution.

2.3 Confirmation of the highest co-expression module via WGCNA

WGCNA is a biological analysis method for calculating the functional sets of various weighted association analyses based on correlation coefficients. In this study, WGCNA was used to obtain the genes in the highest co-expression module. First of all, we used WGCNA to analyze the gene expression information of BC and TC derived from TCGA database. We retained genes with CPM (count per million) ≥ 1 , and processed the missing value. Then the picksoft threshold function was used to filter the soft threshold, so that the network constructed was more accordant with the characteristics of scale-free network. Next the module was segmented by dynamic tree cutting algorithm. According to the correlation coefficient between genes, the hierarchical clustering tree was constructed, and the clustering results were segmented on the basis of the set criteria to obtain different gene modules. Different branches of the cluster tree represented different gene modules with different colors. The correlation between modules and sample traits was analyzed, and the module-trait associations data were visualized using heat map. Finally, the DEGS of BC and TC and the genes in the highest co-expression module obtained by WGCNA analysis of BC and TC were intersected, and the key genes obtained visualized by Venn diagram.

2.4 Functional enrichment analysis

GO annotation can be applied to discover biometric attributes and KEGG pathway enrichment analysis to discover functional attributes[13, 14]. GO Term analysis comprises biological process, cellular component and molecular function. To conduct GO annotation and KEGG pathway enrichment analysis of key genes, we applied cluster profiler package in R to enrich key genes and $P < 0.05$ was taken as the cut-off value.

2.5 PPI Network Construction and acquisition of Hub genes

The STRING database is an online tool for studying PPI that can be used to excavate for hub genes. For the 59 key genes obtained above, we used the String database line tool to search for the interactions between known proteins and predicted proteins. Then, we used cytoscape (version: 3.8.0) to construct a

network model and screened the top 10 genes with the largest MCC value. Moreover, the top 10 genes were defined as core genes.

2.6 Survival analysis and verification of expression patterns of prognostic Genes

The clinical information of BC and TC were gained from TCGA, and patients who completed the follow-up period were chosen for survival analysis with the survival package in R.

Kaplan-Meier (KM) method was performed to estimate the survival rate, and the difference of overall survival (OS) between BC and non-BC, TC and non-TC was compared by Log Rank test. We first performed a prognostic analysis of HUB gene validation, and then the expression patterns of prognostic genes produced in different pathologic tumors and normal tissues were then examined, and $P < 0.05$ was taken as the cut-off value.

2.7 Single-gene GSEA enrichment analysis

Single-gene GSEA was applied to study the influence of a gene expression level on diseases. According to the median expression of the gene to be studied, the tumor samples were grouped into high expression and low expression, and then identify the significantly enriched pathway[15]. For the sake of further probing the biological pathway of prognostic gene pathway involved in the pathogenesis of BC and TC, we performed GSEA. We downloaded c5.all.v7.1.symbols from the GSEA database (<http://www.gsea-msigdb.org/gsea/index.jsp>), grouped them with prognostic genes, and enriched BC and TC tumor samples using the GSEA package in the clusterProfiler package ,and with the cut-off standard of $|\text{normalized enrichment score}| \geq 1.0$, $p \leq 0.05$, $p.\text{adj} \leq 0.25$.

2.8 Statistical Analysis

The gene expression data were all downloaded from the BC and TC databases of TCGA, and all the analyses were performed in R language (Version: 4.0.2). $P < 0.05$ was statistically significant.

3 Results

3.1 SIR for the Second Primary Cancer

As compared to the general population, the SIR of TC in BC patients was 1.2 (95% CI [1.23, 1.26]), and within 2-11 months, the SIR was even higher (2.87, 95% CI [2.52,3.26]), as shown in Table 1. The SIR of BC in TC patients was 1.12 (95% CI [1.07, 1.18]), indicating a certain association between BC and TC.

Table 1
Occurrence of Second Cancers Followings First Primary Breast/Thyroid Cancer

Primary neoplasm site	Second primary neoplasm site	Latency	Observed	Expected	SIR	CI Lower	CI Upper
Breast	Thyroid	2-11 months	242	84.29	2.87#	2.52	3.26
Breast	Thyroid	12-59 months	452	337.9	1.34#	1.22	1.47
Breast	Thyroid	60-119 months	330	300.42	1.1	0.98	1.22
Breast	Thyroid	120+ months	471	432.08	1.09	0.99	1.19
Breast	Thyroid	Total	1,495	1,154.69	1.29#	1.23	1.36
Thyroid	Breast	2-11 months	88	83.25	1.06	0.85	1.3
Thyroid	Breast	12-59 months	418	363.26	1.15#	1.04	1.27
Thyroid	Breast	60-119 months	441	366.57	1.20#	1.09	1.32
Thyroid	Breast	120+ months	916	843.29	1.09#	1.02	1.16
Thyroid	Breast	Total	1,863	1,656.37	1.12#	1.07	1.18
SEER cancer registries, time: 1973-2015; Observed: observed number; Expected: expected number; CI, confidence interval; #: p<0.05; SIR: standardized incidence ratio.							

3.2 DEGs and Co-expression Modules

We obtained the gene expression profiles and clinical information of BC and TC from TCGA database. There were 1222 BC samples, including 1109 tumor tissues and 113 normal tissues. The total number of TC samples was 568, tumor tissue was 510, and normal tissue was 58. The RNA-Seq data for both BC and TC were 56,753 genes.

According to the standard: $|\log FC| \geq 1.0$ and $\text{adj. } P < 0.05$, we use the limma packet for differential expression analysis, and 3736 DEGs (Figure 2A) (Supplementary material 1) were obtained in the BC dataset and 2159 DEGs (Figure 2B) (Supplementary material 2) in the TC dataset.

We used RNA-seq data from TCGA BC database and TC database to construct gene co-expression network, and selected the highest co-expression module. Assign a color to each module, and BC (Figure 2C) produces 11 modules and TC (Figure 2D) produces 6 modules. A heat plot of the module-trait relationships was then used to estimate the correlation between each module and clinical characteristics (cancer and normal). As shown in Figures 2E and 2F, it was the result of the module trait relationship. The turquoise module of BC contained 4149 genes, and the red module of TC contained 162 genes. (turquoise module: $r = 0.75$, $p = 2e-220$; red module: $r = 0.68$, $p = 1e-79$).

3.3 Functional Analysis of the Key Genes

There were 4,149 co-expressed genes in the turquoise module of the BC dataset, and 162 co-expressed genes in the red module of the TC data set. There were 3736 DEGs in the BC data set, and TC has 2159 DEGs. 59 key genes in all were obtained from the intersection of the above four sets, and visualized by Venn plot (Figure 3A).

We used GO annotation (Figure 3B) and KEGG pathway enrichment analysis (Figure 3C) to analyze 59 key genes. In terms of biological process, the analysis result was extracellular matrix organization and extracellular structure organization. For cellular component, the analysis result was collagen-containing extracellular matrix. Finally, in terms of molecular function, they were mostly enriched in the extracellular matrix structural constituent. We also used KEGG pathway enrichment analysis to determine the regulatory pathways. The analysis results shown that the regulated pathways include PI3K-Akt signaling pathway, focal adhesion, and other pathways related to BC and TC.

3.4 Obtaining Prognostic Genes from Hub Genes

We used STRING online data and Cytoscape software to construct 59 key genes Constructed PPI Network and finally obtained the results in (Figure 4A). Then, the 10 largest hub genes of MCC were screened out, which were: DPT, OGN, LAMA2, FBLN5, TNXB, MYH11, EFEMP1, MFAP4, ADH1B, and MYOC (Figure 4B).

We analyzed the survival of 10 hub genes in BC and TC using the R survival package (Supplementary Figure 1, Supplementary Figure 2) and got FBLN5 which were related to OS of BC (Figure 4C) and TC (Figure 4D). Next, Wilcoxon rank sum test (Figure 4E, Figure 4F) and pairing (Figure 4G, Figure 4H) were used to validated the expression level of FBLN5 in BC and TC samples. The results showed that compared with normal tissue, FBLN5 was significantly down-regulated in BC and TC .

3.5 GSEA Results

To determine the exact pathway that FBLN5 regulates in BC and TC, we performed GSEA using RNA-seq data of BC and TC. In the predefined "KEGG pathway "genomes, the main enrichment pathways of BC were CYTOKINE CYTOKINE RECEPTOR INTERACTION, PRIMARY IMMUNODEFICIENCY etc (Figure 5A).

Among the phenotypes with high FBLN5 expression, a total of 41 significant gene sets were found, of which 36 were up-regulated. The main pathways of TC enrichment were: PRIMARY IMMUNODEFICIENCY, STEROID HORMONE BIOSYNTHESIS, COMPLEMENT AND COAGULATION CASCADES etc (Figure 5B). Among the phenotypes with high FBLN5 expression, a total of 47 significant gene sets were found, of which 42 were up-regulated. The common enrichment pathways of the BC and TC mainly included: AUTOIMMUNE THYROID DISEASE, ANTIGEN PROCESSING AND PRESENTATION, B CELL RECEPTOR SIGNALING PATHWAY, PATHWAYS IN CANCER, STEROID HORMONE BIOSYNTHESIS, T CELL RECEPTOR SIGNALING PATHWAY. Most of these pathways were related to immunity, hormones and cancer (Figure 5C, 5D).

4 Discussion

Both the BC and TC have very high morbidity and mortality among the global population, which cause great harm to women and impose a heavy burden on society. The molecular mechanisms underlying the pathogenesis of BC and TC remain unclear. Therefore, it is imperative to identify the potential common genes that contribute to the sequential development of BC and TC.

From the bioinformatics analysis of BC and TC from TCGA, we found 59 key genes with the identical expression trend. These key genes were mostly enriched in the PI3K/Akt signaling pathway. Lots of studies have shown that PI3K/Akt signaling pathways were associated with cell proliferation, differentiation, and apoptosis, and that their component activation is pertinent to tumor occurrence, invasion, and metastasis[16]. PTEN is a lipid phosphatase that targets the PI3K pathway and is frequently mutated and deleted in a variety of cancers, including BC and TC[17]. It was shown that simultaneous inhibition of MAP kinase, PI3K/Akt pathways and HDAC induced certain non-thyroid cancer cells to strongly express sodium iodide (NIS) and ingest radioactive iodine[18]. The mammary gland, like the thyroid gland, has NIS and so may collect iodine as thyroid cells do [19]. Furthermore, NIS symbiosis plays a significant part in the process of iodine from extracellular space to thyroid cells to synthesize thyroid hormone. Fifty-nine key genes were also enriched in these pathways, such as focal adhesion and ECM-receptor interactions. Adhesion kinase is a central regulator of adhesion foci, which affects cell proliferation and migration, and studies addressed overexpression of focal adhesion kinase might be a part of the mechanism of TC invasion and metastasis[20]. ECM protein can stimulate invasive tumor cell programming and thus influence the survival of tumor cells[21-23]. Focal adhesions are mechanically linked to cellular communication centers and ECM cells and are associated with tumor invasion. A variety of studies addressed that the early migration of tumor cells and invasion of adjacent tissues were mediated by focal adhesion signals[24].

In addition, we used String database and Cytoscape software to construct PPI, and defined the 10 largest MCC genes as hub genes. Then we performed prognosis analysis on the 10 hub genes and obtained FBLN5, a gene related to the prognosis of both BC and TC. In the analysis, we found that FBLN5 was

down-regulated in both TC and BC. The FBLN5 gene encodes a protein called Fibulin-5, which is a secreted extracellular matrix protein. FBLN-5 can affect vascular development and transformation[25], with close relation to the development and spread of tumors. Studies suggested that FBLN-5 can prevent tumor angiogenesis and thus inhibit tumor formation[26]. Hiromi et al[27] suggested that Fibulin-5 could affect tumor growth, which might be related to the balance between the antiangiogenic properties of fibulin-5 and its direct effect on tumor cell proliferation and migration. YONG et al[28] suggested that FBLN5 enhanced TGF- β (transforming growth factor- β) and induced EMT (epithelial-to-mesenchymal transition), which are key processes conferring to the occurrence and progression of BC. A series of studies have suggested that integrin-binding domain of FBLN5 mutations can reduce the growth of pancreatic ductal adenocarcinoma[28]. In addition, the down-regulation of Fibulin-5 can affect the prognosis of ovarian cancer, liver cancer and other cancers[29, 30].

Finally, we grouped FBLN5 into GSEA in all tumor samples from BC and TC, respectively. And we found an array of pathways enriched in BC and TC were involved in immune, hormonal, and carcinogenesis. Second primary neoplasm may be caused by physical weakness of the person suffering from cancer, long-term exposure to cancer risk factors, destruction of immune tissues, disruption of body balance and some genetic mutations. There may also be immune deficiency, where long-term exposure to carcinogenic environments makes the body vulnerable to developing cancer due to a lack of immunity to various types of cancer[31].

Therefore, we believe that FBLN5 act a pivotal part in the occurrence and development of BC and TC, but the specific mechanism still needs to be verified with future experimental and clinical evidences.

5 Conclusion

To our knowledge, this is the first co-expression network being constructed using the WGCNA approach to investigate breast- and thyroid-cancer related modules and enrichment pathways. The most co-enriched pathway in BC and TC was PI3K/AKT. In addition, the FBLN5 was the most relevant gene with tumor suppressive potential in both BC and TC, with down-stream pathways involving immunity, hormone biosynthesis and carcinogenesis. Further studies on the specific function of FBLN5 in BC and TC are warrant.

Abbreviations

BC: breast cancer ; TC: thyroid cancer ; SEER: The Surveillance, Epidemiology, and End Results; SIRs: Standardized incidence ratios; TCGA: the cancer genome atlas

WGCNA: Weighted Gene Co-expression Network Analysis ; GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes; CI: confidence interval; DGE: differential gene expression; DEGs: differentially expressed genes; OS: overall survival; FDR: false discovery Rate; FC: fold change; GSEA: gene set enrichment analysis; FBLN5: recombinant fibulin 5

Declarations

Data availability

SIR for the second primary cancer can be obtained: SEER database software (version:8.3.5).

RNA-seq and clinical data for BC and TC can be downloaded: <https://tcga-data.nci.nih.gov/tcga/>

STRING database: <https://string-db.org/>,version,11.0

Acknowledgement

Not applicable.

Authors' contributions

ZMX and DZ conceived and designed the project. ZMX, YSX, CYX and QX conducted the database analysis. ZMX, QX, WDW, ZHW, YKL and HYL analyzed the data and prepared the figures. ZMX and DZ wrote the manuscript. DZ approved the final version to be submitted. All authors read and approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors read and approved the manuscript.

Funding

This work was partly supported by the Natural Science Foundation of Guang

dong Province, China (No. 2019A1515010239); Shantou Municipal health science and technology project (No. 190819145262877); Innovation and Entrepreneurship Training Project for College Students in 2019, China (no. 201810560034 and no. 201810560038); Science and Technology Special Fund of Guangdong Province, China (No. 190829105556145); Strategic and Special Fund for Science and Technology Innovation of Guangdong Province, China (No. 180918114960704); Guangdong Provincial Key Laboratory for Breast Cancer Diagnosis and Treatment (2017B030314116); Interdisciplinary project of Li-Ka-Shing Foundation (No.2020LKSFG05C); Guangdong Basic and Applied Basic Research Foundation (No. 2019A1515110953); Funded by China Postdoctoral Science Foundation (2020M672753); Supported by Department of Education of Guangdong Province (2019KQNCX035, 2019KQNCX033). Availability of data and materials. All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethical approval

Not applicable.

Consent for publication

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**(6):394-424.
2. An JH, Hwangbo Y, Ahn HY, Keam B, Lee KE, Han W, Park DJ, Park IA, Noh DY, Youn YK *et al*: **A Possible Association Between Thyroid Cancer and Breast Cancer.** *THYROID* 2015, **25**(12):1330-1338.
3. Kim YA, Kim YA, Cho SW, Song YS, Min HS, Park IA, Park DJ, Hwang KT, Park YJ: **Increased expression of thyroid hormone receptor alpha and estrogen receptor alpha in breast cancer associated with thyroid cancer.** *Eur J Surg Oncol* 2021.
4. Huang NS, Chen XX, Wei WJ, Mo M, Chen JY, Ma B, Yang SW, Xu WB, Wu J, Ji QH *et al*: **Association between breast cancer and thyroid cancer: A study based on 13 978 patients with breast cancer.** *Cancer Med* 2018, **7**(12):6393-6400.
5. Li S, Yang J, Shen Y, Zhao X, Zhang L, Wang B, Li P, Wang Y, Yi M, Yang J: **Clinicopathological features, survival and risk in breast cancer survivors with thyroid cancer: an analysis of the SEER database.** *BMC PUBLIC HEALTH* 2019, **19**(1):1592.
6. Nielsen SM, White MG, Hong S, Aschebrook-Kilfoy B, Kaplan EL, Angelos P, Kulkarni SA, Olopade OI, Grogan RH: **The Breast-Thyroid Cancer Link: A Systematic Review and Meta-analysis.** *Cancer Epidemiol Biomarkers Prev* 2016, **25**(2):231-238.
7. Can T: **Introduction to bioinformatics.** *Methods Mol Biol* 2014, **1107**:51-71.
8. Segundo-Val IS, Sanz-Lozano CS: **Introduction to the Gene Expression Analysis.** *Methods Mol Biol* 2016, **1434**:29-43.
9. Zhang B, Horvath S: **A general framework for weighted gene co-expression network analysis.** *Stat Appl Genet Mol Biol* 2005, **4**:e17.
10. Li Z, Cui Y, Feng J, Guo Y: **Identifying the pattern of immune related cells and genes in the peripheral blood of ischemic stroke.** *J TRANSL MED* 2020, **18**(1):296.
11. Ren Y, Li W, Liu S, Li Z, Wang J, Yang H, Xu Y: **A Weighted Gene Co-expression Network Analysis Reveals lncRNA Abnormalities in the Peripheral Blood Associated With Ultra-High-Risk for Psychosis.** *FRONT PSYCHIATRY* 2020, **11**:580307.

12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *NUCLEIC ACIDS RES* 2015, **43**(7):e47.
13. **The Gene Ontology (GO) project in 2006.** *NUCLEIC ACIDS RES* 2006, **34**(Database issue):D322-D326.
14. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *NUCLEIC ACIDS RES* 2000, **28**(1):27-30.
15. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al.* **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**(43):15545-15550.
16. Verret B, Cortes J, Bachelot T, Andre F, Arnedos M: **Efficacy of PI3K inhibitors in advanced breast cancer.** *ANN ONCOL* 2019, **30**(Suppl_10):x12-x20.
17. Csolle MP, Ooms LM, Papa A, Mitchell CA: **PTEN and Other PtdIns(3,4,5)P3 Lipid Phosphatases in Breast Cancer.** *INT J MOL SCI* 2020, **21**(23).
18. Liu Z, Xing M: **Induction of sodium/iodide symporter (NIS) expression and radioiodine uptake in non-thyroid cancer cells.** *PLOS ONE* 2012, **7**(2):e31729.
19. Consorti F, Di Tanna G, Milazzo F, Antonaci A: **Nulliparity enhances the risk of second primary malignancy of the breast in a cohort of women treated for thyroid cancer.** *WORLD J SURG ONCOL* 2011, **9**:88.
20. Owens LV, Xu L, Dent GA, Yang X, Sturge GC, Craven RJ, Cance WG: **Focal adhesion kinase as a marker of invasive potential in differentiated human thyroid cancer.** *ANN SURG ONCOL* 1996, **3**(1):100-105.
21. Aguilera KY, Rivera LB, Hur H, Carbon JG, Toombs JE, Goldstein CD, Dellinger MT, Castrillon DH, Brekken RA: **Collagen signaling enhances tumor progression after anti-VEGF therapy in a murine model of pancreatic ductal adenocarcinoma.** *CANCER RES* 2014, **74**(4):1032-1044.
22. Sethi T, Rintoul RC, Moore SM, MacKinnon AC, Salter D, Choo C, Chilvers ER, Dransfield I, Donnelly SC, Strieter R *et al.* **Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: a mechanism for small cell lung cancer growth and drug resistance in vivo.** *NAT MED* 1999, **5**(6):662-668.
23. Meredith JJ, Fazeli B, Schwartz MA: **The extracellular matrix as a cell survival factor.** *MOL BIOL CELL* 1993, **4**(9):953-961.
24. Eke I, Cordes N: **Focal adhesion signaling and therapy resistance in cancer.** *SEMIN CANCER BIOL* 2015, **31**:65-75.
25. Nakamura T, Ruiz-Lozano P, Lindner V, Yabe D, Taniwaki M, Furukawa Y, Kobuke K, Tashiro K, Lu Z, Andon NL *et al.* **DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries.** *J BIOL CHEM* 1999, **274**(32):22476-22483.
26. Albig AR, Neil JR, Schiemann WP: **Fibulins 3 and 5 antagonize tumor angiogenesis in vivo.** *CANCER RES* 2006, **66**(5):2621-2629.

27. Yanagisawa H, Schluterman MK, Brekken RA: **Fibulin-5, an integrin-binding matricellular protein: its function in development and disease.** *J CELL COMMUN SIGNAL* 2009, **3**(3-4):337-347.
28. Lee YH, Albig AR, Regner M, Schiemann BJ, Schiemann WP: **Fibulin-5 initiates epithelial-mesenchymal transition (EMT) and enhances EMT induced by TGF-beta in mammary epithelial cells via a MMP-dependent mechanism.** *CARCINOGENESIS* 2008, **29**(12):2243-2251.
29. Li R, Wu H, Jiang H, Wang Q, Dou Z, Ma H, Yan S, Yuan C, Yang N, Kong B: **FBLN5 is targeted by microRNA27a3p and suppresses tumorigenesis and progression in highgrade serous ovarian carcinoma.** *ONCOL REP* 2020, **44**(5):2143-2151.
30. He J, Zhao H, Deng D, Wang Y, Zhang X, Zhao H, Xu Z: **Screening of significant biomarkers related with prognosis of liver cancer by lncRNA-associated ceRNAs analysis.** *J CELL PHYSIOL* 2020, **235**(3):2464-2477.
31. Bolf EL, Sprague BL, Carr FE: **A Linkage Between Thyroid and Breast Cancer: A Common Etiology?** *Cancer Epidemiol Biomarkers Prev* 2019, **28**(4):643-649.

Figures

Figure 1

The research workflow of this study.

Figure 2

(A) Volcano plot of DEGs in BC; (B) Volcano plot of DEGs in TC. (C) The cluster dendrogram of co-expression network modules of BC each module was allocated different colors. (D) The cluster dendrogram of co-expression network modules of TC, each module was allocated different colors. (E) Module-trait relationships of BC, a unit contained the corresponding correlation and P values. (F) Module-trait relationships of TC, a unit contained the corresponding correlation and P values.

Figure 3

(A) Venn diagram showed 59 overlapping genes in the DEG lists of BC and TC, and co-expression modules. (B) GO enrichment analysis for 59 overlapping genes. (C) KEGG enrichment analysis for 59 overlapping genes.

Figure 4

(A) Construction of PPI and visualization of PPI networks; (B) Obtain the top 10 hub genes with the largest MCC value. (C) OS analysis of FBLN5 in BC patients. (D) OS analysis of FBLN5 in TC patients. (E) Wilcoxon rank method was used to test the expression level of FBLN5 in BC and normal tissues. (F) Wilcoxon rank method was used to test the expression level of FBLN5 in TC and normal tissues. (G) Paired differentiation analysis for expression of FBLN5 in the normal and BC sample deriving from the same one patient. (H) Paired differentiation analysis for expression of FBLN5 in the normal and TC sample deriving from the same one patient.

Figure 5

GSEA for samples with FBLN5 expression. (A) Group all tumor samples in BC with FBLN5 for KEGG: show the top 5 significant enrich pathways;(B) Group all tumor samples in TC with FBLN5 for KEGG: show the top 5 significant enrich pathways; (C) Group all tumor samples in BC with FBLN5 for KEGG: show the significant enrich pathways; (D) Group all tumor samples in TC with FBLN5 for KEGG: show the significant enrich pathways.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFile.doc](#)