

Article

# Network Analyses of Integrated Differentially Expressed Genes in Papillary Thyroid Carcinoma to Identify Characteristic Genes

Junliang Shang <sup>1,2</sup> , Qian Ding <sup>2</sup>, Shasha Yuan <sup>2</sup>, Jin-Xing Liu <sup>2</sup>, Feng Li <sup>3</sup>  and Honghai Zhang <sup>4,\*</sup>

<sup>1</sup> School of Statistics, Qufu Normal University, Qufu 273165, China; shangjunliang110@163.com

<sup>2</sup> School of Information Science and Engineering, Qufu Normal University, Rizhao 276800, China; dingqian19@126.com (Q.D.); jiayouyiss@126.com (S.Y.); sdcavell@126.com (J.-X.L.)

<sup>3</sup> School of Computer Science and Technology, Xidian University, Xi'an 710071, China; lifeng\_10\_28@163.com

<sup>4</sup> College of Life Science, Qufu Normal University, Qufu 273165, China

\* Correspondence: zhanghonghai67@126.com; Tel.: +86-63-3398-1244

Received: 8 November 2018; Accepted: 9 January 2019; Published: 14 January 2019



**Abstract:** Papillary thyroid carcinoma (PTC) is the most common type of thyroid cancer. Identifying characteristic genes of PTC are of great importance to reveal its potential genetic mechanisms. In this paper, we proposed a framework, as well as a measure named Normalized Centrality Measure (NCM), to identify characteristic genes of PTC. The framework consisted of four steps. First, both up-regulated genes and down-regulated genes, collectively called differentially expressed genes (DEGs), were screened and integrated together from four datasets, that is, GSE3467, GSE3678, GSE33630, and GSE58545; second, an interaction network of DEGs was constructed, where each node represented a gene and each edge represented an interaction between linking nodes; third, both traditional measures and the NCM measure were used to analyze the topological properties of each node in the network. Compared with traditional measures, more genes related to PTC were identified by the NCM measure; fourth, by mining the high-density subgraphs of this network and performing Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, several meaningful results were captured, most of which were demonstrated to be associated with PTC. The experimental results proved that this network framework and the NCM measure are useful for identifying more characteristic genes of PTC.

**Keywords:** differentially expressed genes (DEGs); interaction network; papillary thyroid cancer (PTC); Normalized Centrality Measure (NCM)

## 1. Introduction

The morbidity rate of papillary thyroid carcinoma (PTC) ranks the highest among thyroid cancer. Exploring the cause and underlying molecular events of PTC has become a hot topic [1], and various methods for analyzing characteristic genes of PTC have emerged. There are many methods for exploring the characteristic genes of PTC. Kim et al. [2] proposed to select differentially expressed genes (DEGs) by using paired *t*-test, linear models for microarray data, and significance analysis of microarrays. Chevillard et al. [3] used microarrays containing probes to hybridize with RNA, which came from normal and tumor thyroid samples labeled with Cy3-dUTP and Cy5-dUTP, respectively. Jarzab et al. [4] proposed a potent molecular classifier able to discriminate between PTC and nonmalignant thyroid, which used a support vector machines-based technique. However, most of the existing methods focus on analyzing differences between genes in two experimental conditions or using statistical methods to analyze interactions between functionally similar genes. It becomes

increasingly difficult to identify the characteristic genes from analysis results with the diversity and high dimensionality of biosequencing data, which makes the advantages of network mining methods more obvious.

The methods of network mining can systematically reflect the interactions between biomolecules; thus, it is beneficial for researchers to fully understand how various biomolecules in biological cells interacting and performing biological function processes. He et al. [5] proposed a method to identify molecular markers related to the clinicopathological features of renal cell carcinoma through a weighted gene co-expression network, but multiple thresholds need to be set through the hub gene of the Weighted Gene Co-expression Network Analysis (WGCNA) calculation module. Zhang et al. [6] proposed a method for analyzing gene interaction enrichment, which combines the knowledge of gene pathways to identify interaction effects on enriched genes. This method ignores genes that have no direct effect on expression, but interact with other genes to affect expression. Kamran et al. [7] proposed a new measure of the topology properties of the network; however, this measure ignores the relationship between the various factors of the nodes in the formula and the importance of the nodes. Wei et al. [8] proposed the Length-Net-Driver (LNDriver) method to construct a gene-gene interaction network, which integrates mutation data and expression data into the network, and finally uses the greedy algorithm to select driver genes. The genes obtained by this method only considering the local features and do not consider the role in the global network. In addition, existing network methods to mine characteristic genes are mostly limited to considering only local features or global features.

In this paper, we propose a framework to identify the characteristic genes of PTC. First, DEGs identified from multiple datasets were integrated and used to build an interaction network. The nodes in the network represent genes, the edges between two nodes indicating that these two nodes interact with each other. The weight of an edge was quantified by a comprehensive scoring mechanism, which was formed by comprehensive data, data mined from PubMed abstract text, database data, and results predicted using bioinformatics methods. Next, in order to identify the characteristic genes, a new measure Normalized Centrality Measure (NCM) is proposed in this paper, which takes into account both local and global features of nodes. Compared with the existing measures, characteristic genes identified by NCM were more biologically significant. Moreover, we used the previously proposed Hypergraph Construction and High-Density Subgraph Detection (HC-HDSD) method and performed Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to supplement the experimental results. Our research framework and experimental results provide reliable molecular markers for early detection and prognosis.

## 2. Methods

### 2.1. Datasets

Four profile datasets, i.e., GSE3467 [9], GSE3678 [10], GSE33630 [11], and GSE58545 [12], were downloaded from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) for this study. Platforms, sample numbers, and citations of these four datasets are shown in Table 1.

**Table 1.** Samples distribution of papillary thyroid carcinoma (PTC) dataset.

GEO	Platform	Normal	Tumor	Reference
GSE3467	GPL570	9	9	He et al. [9]
GSE3678	GPL570	7	7	He et al. [10]
GSE33630	GPL570	49	45	Tomás et al. [11]
GSE58545	GPL96	27	18	Rusinek et al. [12]

### 2.2. Identification of Differentially Expressed Genes

Raw data were converted into an expression matrix and the *t*-test method in limma package [13] was subsequently used to identify DEGs between PTC tissue samples and normal thyroids [14]. In the

expression matrix, one row represents one sample and one column represents one gene. A gene which satisfied the conditions of  $p$ -value (adjusted by Benjamini & Hochberg)  $< 0.05$  and  $\log_{2}FC > 2$  (FC, fold change) was identified as an up-regulated gene. A gene which satisfied conditions of  $p$ -value (adjusted by Benjamini & Hochberg)  $< 0.05$  and  $\log_{2}FC < -2$  was identified as a down-regulated gene. All identified up-regulated genes and down-regulated genes were integrated into DEGs. Description of  $p$ -value and  $\log_{2}FC$  can be seen in [13].

### 2.3. Construction of Interaction Network of Differentially Expressed Genes

The identified DGEs were utilized to construct a gene-gene interaction network. The nodes represented genes, the tag names of nodes correspond to gene names, and the edges between nodes indicate interactions between the two nodes. Sources of network edges contain experimental data, data mined from PubMed abstract text, database data, and results predicted using gene contiguous, gene fusion, phylogenetic profile, and gene co-expression based on chip data and other bioinformatics methods, which were based on the STRING database (<http://string-db.org>) [15]. The composite score, which was systematically calculated by different methods described above, were used as weights of edges in the constructed network. Isolated nodes that were not connected to other nodes were deleted. In addition, the size of a node indicated the number of neighbors the node had, that is, the larger the node, the more edges that were connected to this node.

### 2.4. Analyses of Differentially Expressed Genes Interaction Network

#### 2.4.1. Traditional Topological Properties

There were many measures to quantify the topology properties, currently. The following traditional five measures, which we selected, are most widely used to quantify the topological properties of a network. The definitions of them are as follows.

The Degree Centrality (DC) of node  $v$  represents the number of nodes which are directly connected to  $v$  in the network. It is defined as:

$$DC(v) = \sum_{j \in G} \frac{a_{vj}}{n-1} \quad (1)$$

where  $n$  is the total nodes in the network and  $a_{vj}$  represents that node  $v$  and node  $j$  are directly connected. DC is the most basic topological property to reflect local features of the node in the network.

Betweenness Centrality (BC) is a measure to quantify the number of times the node  $v$  appearing in the shortest path between other nodes, which is defined as:

$$BC(v) = \sum_{i \neq j \neq v \in V} \frac{\sigma_{ivj}}{\sigma_{ij}} \quad (2)$$

where  $\sigma_{i,j}$  represents the number of the shortest paths from node  $i$  to node  $j$ ,  $\sigma_{ivj}$  represents the number of paths through node  $v$ . BC reflects the global features of the node. The higher the value of BC, the more important the node is in keeping the network tightly connected.

Closeness Centrality (CC) is a measure to quantify the average shortest distance from one node to all other nodes in the network. The CC of node  $v$  is defined as follows:

$$CC(v) = \frac{1}{n-1} \sum_{j \neq v \in V} d_{vj} \quad (3)$$

where  $n$  is the number of nodes,  $d_{vj}$  represents the shortest distance from node  $v$  to node  $j$ . The CC measure reflects the extent of node  $v$  closing to the "center" of the network. The smaller the value of CC, the closer the node is to the center, which represents the global features of the node.

Clustering Coefficient (CLC) embodies the dense connection nature between some nodes, which is defined as:

$$CLC(v) = \frac{2n}{k(k-1)} \quad (4)$$

where  $k$  represents the number of neighbors of node  $v$ , and  $n$  represents the number of sides connected between  $k$  neighbors of node  $v$ . A higher value of CLC indicates that the node is more closely connected to its neighbors.

The Eccentricity Centrality (EC) is the maximum of the distance from one node to others in the network. The eccentricity of node  $v$  in the network  $G$  is defined as:

$$EC_G(v) = \frac{1}{\max_{V \in G} \{dist(v, u)\}} \quad (5)$$

where  $u$  is a node in the network  $G$ .  $dist(v, u)$  is the largest geodesic distance between the node  $v$  and other nodes. The smaller the value of EC, the more important the node is in the network.

#### 2.4.2. Normalized Centrality Measure

Considering that almost all the existing measures consider only one of the local features and the global features, the results obtained may be relatively one-sided. Therefore, in this paper, we propose a new measure, Normalized Centrality Measure (NCM), by combining the three traditional measures: degree centrality, betweenness centrality, and closeness centrality to identify important genes by calculating the value of NCM for node  $v$ :

$$NCM(v) = \left( \frac{BC(v)}{BC_{\max}} + \frac{DC(v)}{DC_{\max}} \right) / \frac{CC(v)}{CC_{\max}} \quad (6)$$

where  $DC(v)$  is the degree centrality of node  $v$  and  $DC_{\max}$  is the maximum value of degree centrality in the network;  $CC(v)$  is the closeness centrality of node  $v$  and  $CC_{\max}$  is the maximum value of closeness centrality in the network;  $BC(v)$  is the betweenness centrality of node  $v$  and  $BC_{\max}$  is the maximum value of betweenness centrality in the network. NCM comprehensively reflects both the global and local features of nodes. The larger the value of the  $NCM(v)$  of the node  $v$ , the more important the node is. From a theoretical analysis, this measure is meaningful.

#### 2.4.3. Detection of High-Density Subgraphs

In addition to calculating the topological properties of each node, this paper also introduces the HC-HDSD method to extract high-density subgraphs, which was proposed in our previous work. The high-density subgraph was actually the interactions between several genes that may cause PTC, which is in the form of cluster.

First, the Maximal Clique Centrality (MCC) method [16] to detect the hub nodes and their subgraphs:

$$MCC(v) = \sum_{C \in S(v)} (|C|-1)! \quad (7)$$

where  $S(v)$  is the set contains the maximal clique group of  $v$ ,  $C$  is a subset of  $S(v)$ ,  $(|C|-1)!$  is the product of all positive integers, which were less than the number of nodes in the maximal clique  $C$ .

Then, based on the subgraphs obtained by the above method, the NCMine method [17] is used to further detect meaningful high-density subgraphs. The threshold cliqueness used to determine whether to extract the high-density subgraph is defined as follows:

$$cliqueness(G) = \frac{E(G)}{C_{V(G)}^2} \quad (8)$$

where  $G$  is the new module after adding a node,  $E(G)$  is the number of edges of the new module, and  $V(G)$  is the number of nodes of the new module.

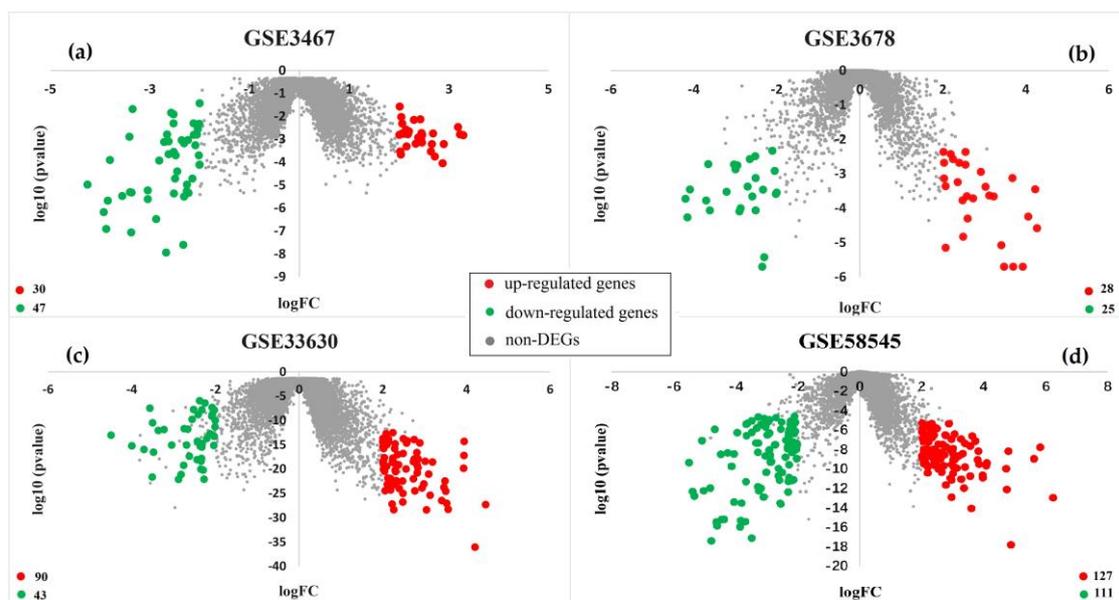
### 2.5. Gene Ontology and Pathway Enrichment Analysis

Gene ontology analysis is used to annotate genes as well as gene products, and identify biological characteristic for DEGs. GO annotations are performed on the DEGs by using a DAVID online tool (<https://david.ncifcrf.gov/>) [18], which was an essential foundation for the success of high-throughput gene function analysis. In addition, KEGG pathway analysis of DEGs is performed using the KOBAS online analysis database (version 3.0) (<http://kobas.cbi.pku.edu.cn/>) [19]. In this paper, we analyzed the DEGs, which were derived from the integrated up-regulated and down-regulated genes from four different microarray datasets;  $p$ -value < 0.05 was set as the cut-off criterion.

## 3. Results and Discussions

### 3.1. Identification of Differentially Expressed Genes

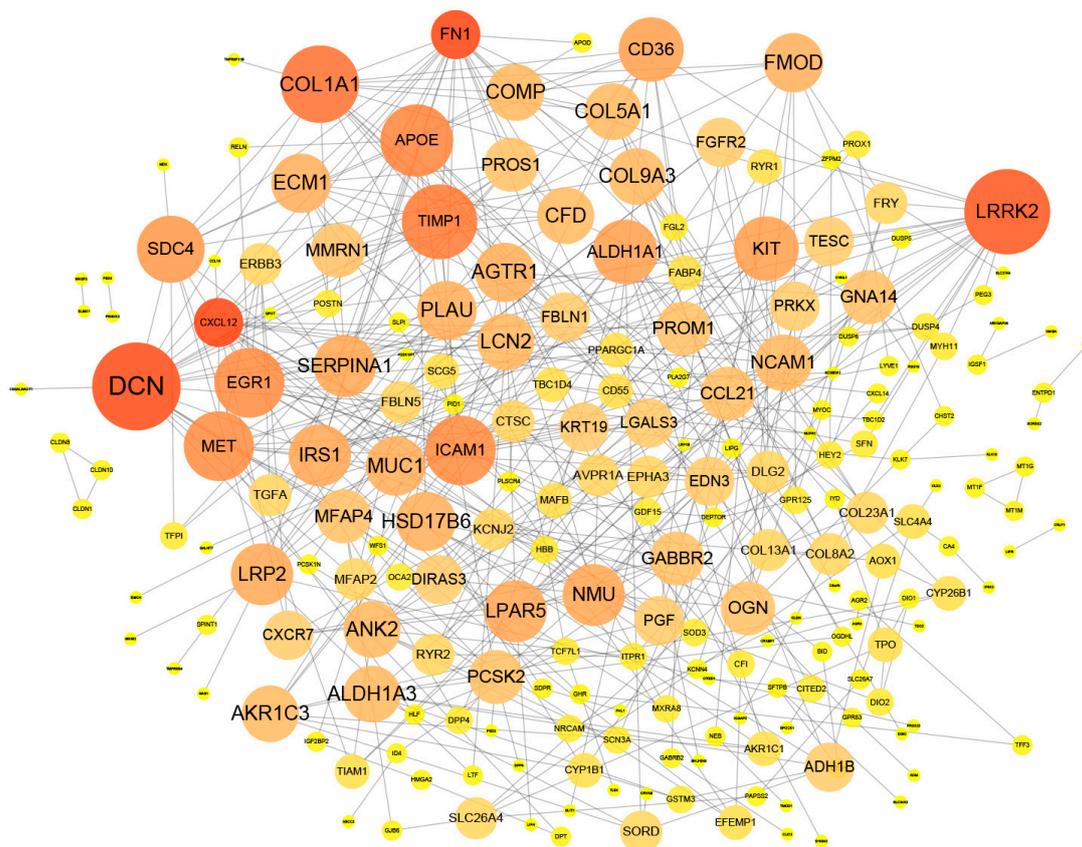
There are 77, 53, 123, 238 DEGs obtained by the limma package from the above four datasets, among which there were 30, 28, 90, and 127 up-regulated genes, as well as 47, 25, 43, and 111 down-regulated genes, respectively. The distribution of DEGs is shown in the following volcano map. The red circles represent up-regulated genes, the green circles represent down-regulated genes, and the gray circles represent non-DEGs. Figure 1a–d show DEGs identified from datasets GSE3467, GSE3678, GSE33630, GSE58545, respectively.



**Figure 1.** Volcano map showing differentially expressed genes (DEGs). (a–d) show DEGs of dataset GSE3467, GSE3678, GSE33630, and GSE58545, respectively.

### 3.2. Analyses of Differentially Expressed Genes Interaction Network

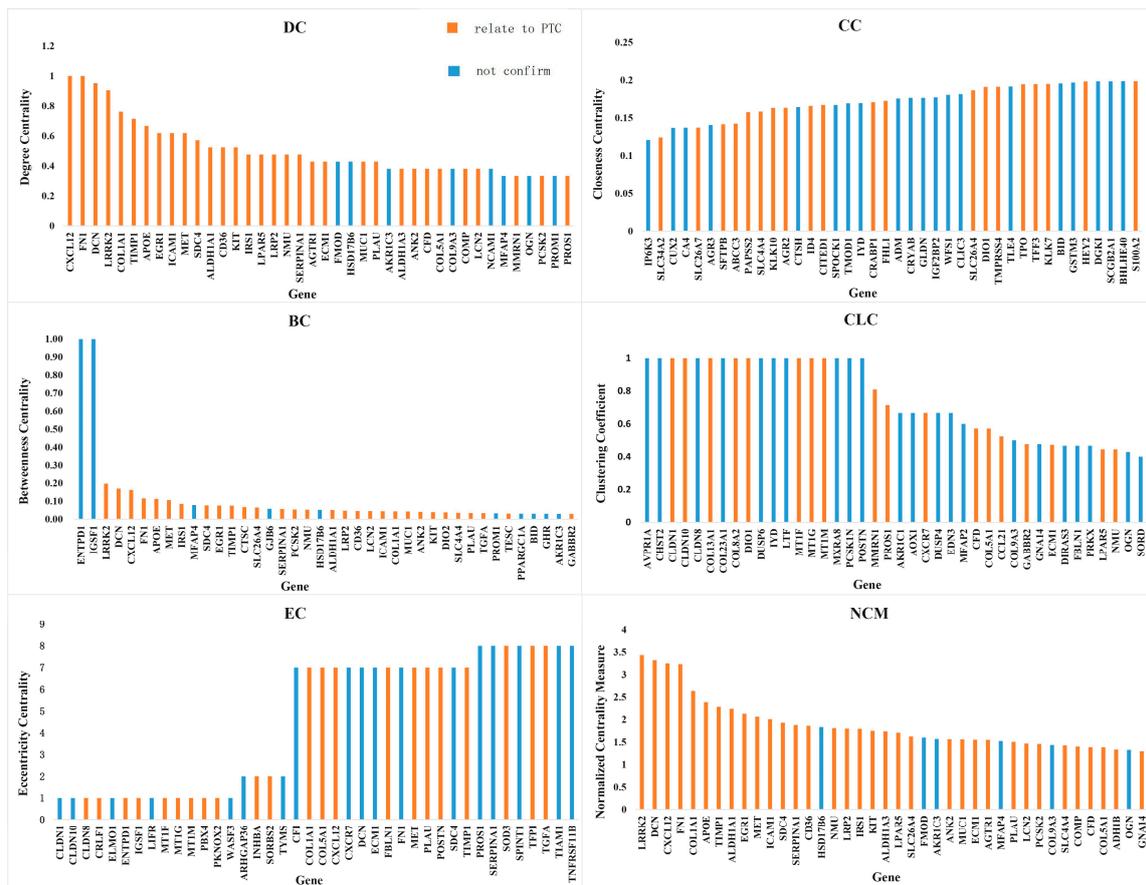
A complex interaction network is obtained from the STRING database to describe interactions between genes, with a total of 501 DEGs including 275 up-regulated genes and 226 down-regulated genes. The nodes in the network represent genes. After removing the duplicate values and the individual nodes that do not connect to other nodes, there were a total of 261 nodes in this network. Otherwise, edges between two nodes represent that these two nodes are interacted; the tags on nodes represent the gene names. The size of the node represents the number of neighbors directly connected to the node. The larger the node, the more neighbors connected to this node. A gene-gene interaction network is established for all identified DEGs, which is shown in the Figure 2.



**Figure 2.** Interaction network of DEGs. The size and color of the node represent the importance of the node. The larger the node and the redder the color, the more neighbors are connected to the node.

### 3.3. Analyses of Characteristic Genes

The topological properties DC, CC, BC, CLC, EC, and NCM of each node in the above network were calculated. The larger the values of DC, BC, CLC and NCM, the more important the node was; the smaller the values of CC and EC, the more important the role that this node played in the network. The values of DC, BC, CLC and NCM were sorted from large to small; the values of CC and EC were sorted from small to large. The relationship between the top 40 values of these topological properties and their corresponding genes is shown in Figure 3. Figure 3a–e show the values of traditional topological property we calculated and their corresponding genes. Figure 3f shows the top 40 genes calculated using our newly proposed measure NCM; among them, there are 34 genes that have been confirmed in the existing literature to be related to the production of PTC. The accuracy was 85%. The values of the logFC and *p*-value of these 40 genes are shown in the tables in the attached file, and 31, 30, 20, 18, and 22 genes were found by calculating the traditional measure of topological properties DC, BC, CC, CLC, and EC. The accuracies were 77.5%, 75%, 50%, 45%, and 55%, respectively. Table 2 shows genes that have been confirmed to be involved in the production of PTC. These genes belong to the top 40 gene calculated by using DC, BC, CC, CLC, EC, and NCM measures. In Table 2, genes with bold font have been recorded in GeneCard database (<http://www.genecards.org/>), of which was confirmed that they are associated with PTC. In addition, other genes have been confirmed to be related to PTC by referring to the existing literature.



**Figure 3.** Histograms of topological property values of genes. The genes represented by the histograms of orange have been confirmed to be related to papillary thyroid carcinoma (PTC); the genes represented by the histograms of blue have not been confirmed to be related to PTC.

Analyses of top 40 genes calculating by the newly proposed topological property NCM measure were as follows. Qiu et al. [20] suggested that *FN1*, *ICAM1* and *TIMP1* play an important roles in the progression of PTC; *COL1A1* [21], *COL5A1* [21], and *PLAU* [21] were confirmed to provide therapeutic biomarkers of PTC. *MET* [22], and *SERPINA1* [22] were screened as common DEGs, which may be related to PTC. *LCN2* [23] transcript levels in tumor tissues were significantly higher in PTC compared with normal tissues. *SDC4* [2] and *LRRK2* [2] were validated that they have a significant genetic difference with normal tissue. *ECM1* [24] and *KIT* [24] were identified as DEGs that contain valuable candidate genes to differentiate related tumor types on the molecular level. Hyun Sook et al. [25] confirmed that the expression of *SLC34A2* and *COMP* increase in PTC compared with normal tissue. *CXCL12* and *IRS1* were identified as central genes in [26], they may be key genes for PTC. *CFD* [27] may be helpful for the development of gene panel for PTC diagnosis. *PCSK2* [28] could be used to identify the benign and malignant follicular thyroid neoplasia. *MUC1* [29] was demonstrated to be an important oncogene in PTC and make an important role in therapeutic cancer vaccine development. *ALDH1A1* [30] was down-regulated in PTC tissues. *SLC26A4* [30] and other genes revealed follicular variant of PTC *CD36* [31] was reported to correlate with neovascularization of PTC *DCN* [32] could help for thyroid tumors to better define a tumor signature and be useful for thyroid tumor treatment. *LRP2* [33] makes effects for thyroid through expressing retained intracellularly. *AGTR1* [34] methylation signatures which involved in the logistic regression model, may related to thyroid cancer. *APOE* [35] was identified as hub gene which may be the predictive risk marker of PTC in the process of progression and prognosis. *NMU* and *LPAR5* may be the targets for thyroid cancer treatment, which are abnormally expressed in thyroid cancer tissues [36]. *ADH1B* [37] could be the

target autoantigens in thyroid cancer and play potential clinical role. Kai-Chun et al. [37] demonstrated that *GNA14* involves in the regulation of the thyroid function and autoimmunity. *ALDH1A3* [37] may be the factor of mediating cancer stemness in thyroid cancer. *EGR1* [38] may have a marked effect on thyroid cancer. Stein [39] identified that *ANK2* gene shows a high expression level in PTC.

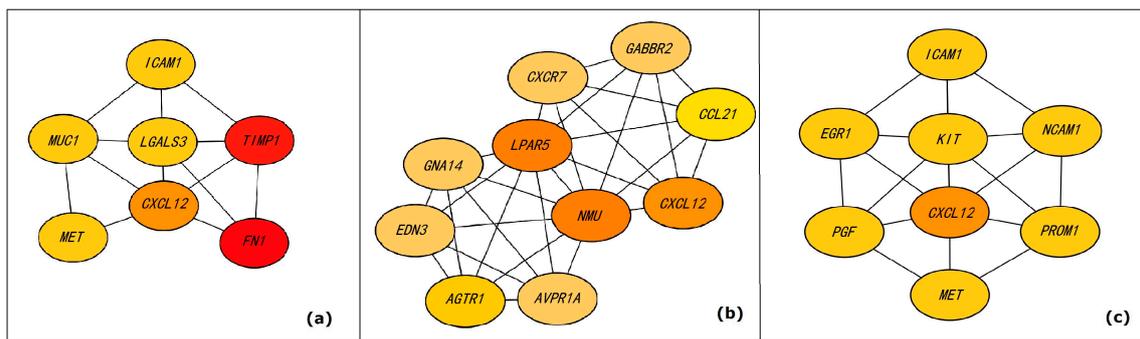
**Table 2.** References of PTC-related genes in top 40 captured genes.

Measures	Count <sup>1</sup>	Genes <sup>2</sup>
Degree Centrality	32	<i>AGTR1</i> [34] <i>ALDH1A1</i> [30] <i>COL1A1</i> [21] <i>COMP</i> [25] <i>CXCL12</i> [26] <i>DCN</i> [32] <i>ECM1</i> [24] <i>FN1</i> [20] <i>ICAM1</i> [20] <i>IRS1</i> [26] <i>KIT</i> [24] <i>LCN2</i> [23] <i>LRP2</i> [33] <i>LRRK2</i> [2] <i>MET</i> [21] <i>MUC1</i> [29] <i>PLAU</i> [20] <i>SDC4</i> [2] <i>SERPINA1</i> [21] <i>TIMP1</i> [20] <i>CD36</i> [31] <i>EGR1</i> [38] <i>ALDH1A3</i> [37] <i>NMU</i> [36] <i>ANK2</i> [39] <i>APOE</i> [35] <i>CFD</i> [27] <i>COL5A1</i> [20] <i>LPAR5</i> [36] <i>PCSK2</i> [28] <i>PROS1</i> [2] <i>MMRN1</i> [27]
Closeness Centrality	20	<i>AGR2</i> [40] <i>CITED1</i> [2] <i>CRABP1</i> [41] <i>KLK10</i> [42] <i>KLK7</i> [25] <i>S100A2</i> [43] <i>SFTPB</i> [44] <i>SLC26A4</i> [30] <i>SLC34A2</i> [25] <i>TFE3</i> [28] <i>TMPPSS4</i> [45] <i>TPO</i> [41] <i>DIO1</i> [32] <i>ID4</i> [45] <i>PAPSS2</i> [39] <i>SLC4A4</i> [25] <i>ABCC3</i> [46] <i>FHL1</i> [38] <i>HEY2</i> [46] <i>SLC26A7</i> [47]
Betweenness Centrality	30	<i>ALDH1A1</i> [30] <i>COL1A1</i> [21] <i>CTSC</i> [48] <i>CXCL12</i> [26] <i>DCN</i> [32] <i>FN1</i> [20] <i>ICAM1</i> [20] <i>IRS1</i> [26] <i>KIT</i> [24] <i>LCN2</i> [23] <i>LRP2</i> [33] <i>LRRK2</i> [2] <i>MET</i> [22] <i>MUC1</i> [29] <i>PLAU</i> [21] <i>SDC4</i> [2] <i>SERPINA1</i> [22] <i>SLC26A4</i> [30] <i>TGFA</i> [41] <i>TIMP1</i> [19] <i>DIO2</i> [32] <i>CD36</i> [31] <i>ANK2</i> [39] <i>APOE</i> [35] <i>GABBR2</i> [24] <i>PCSK2</i> [28] <i>SLC4A4</i> [25] <i>TESC</i> [39] <i>EGR1</i> [38] <i>NMU</i> [36]
Clustering Coefficient	18	<i>CCL21</i> [49] <i>CLDN1</i> [20] <i>CLDN10</i> [50] <i>ECM1</i> [24] <i>DIO1</i> [32] <i>MT1G</i> [51] <i>CFD</i> [27] <i>COL13A1</i> [22] <i>COL5A1</i> [21] <i>CXCR7</i> [52] <i>GABBR2</i> [24] <i>MMRN1</i> [27] <i>PROS1</i> [2] <i>COL8A2</i> [53] <i>LPAR5</i> [36] <i>MT1F</i> [54] <i>MT1M</i> [54] <i>NMU</i> [36]
Eccentricity Centrality	22	<i>CLDN1</i> [20] <i>CLDN10</i> [50] <i>MT1G</i> [51] <i>COL1A1</i> [20] <i>CXCL12</i> [26] <i>DCN</i> [32] <i>ECM1</i> [24] <i>FN1</i> [20] <i>MET</i> [22] <i>PLAU</i> [20] <i>SDC4</i> [2] <i>TIMP1</i> [20] <i>SERPINA1</i> [22] <i>TGFA</i> [41] <i>TNFRSF11B</i> [22] <i>ELMO1</i> [55] <i>IGSF1</i> [55] <i>MT1F</i> [54] <i>SORBS2</i> [39,48] <i>COL5A1</i> [21] <i>CXCR7</i> [52] <i>PROS1</i> [2]
Normalized Centrality Measure	34	<i>ALDH1A1</i> [30] <i>CD36</i> [31] <i>COL1A1</i> [20] <i>COMP</i> [25] <i>CXCL12</i> [26] <i>DCN</i> [32] <i>ECM1</i> [24] <i>FN1</i> [20] <i>ICAM1</i> [20] <i>IRS1</i> [26] <i>KIT</i> [24] <i>LCN2</i> [23] <i>LRP2</i> [33] <i>LRRK2</i> [2] <i>MET</i> [22] <i>MUC1</i> [29] <i>PLAU</i> [21] <i>SDC4</i> [2] <i>SERPINA1</i> [22] <i>SLC26A4</i> [30] <i>TIMP1</i> [20] <i>AGTR1</i> [34] <i>ANK2</i> [39] <i>APOE</i> [35] <i>CFD</i> [27] <i>COL5A1</i> [21] <i>LPAR5</i> [36] <i>PCSK2</i> [28] <i>SLC4A4</i> [25] <i>ADH1B</i> [37] <i>GNA14</i> [37] <i>NMU</i> [36] <i>ALDH1A3</i> [37] <i>EGR1</i> [38]

<sup>1</sup> Count: The number of genes that have been confirmed to be related to PTC from top 40 genes obtained by six measures. <sup>2</sup> Genes: Details of existing literature that described these genes to be involved in the production of PTC. DC: Degree Centrality; CC: Closeness Centrality; BC: Betweenness Centrality; CLC: Clustering Coefficient; EC: Eccentricity Centrality; NCM: Normalized Centrality Measure.

### 3.4. Analyses of High-Density Subgraphs

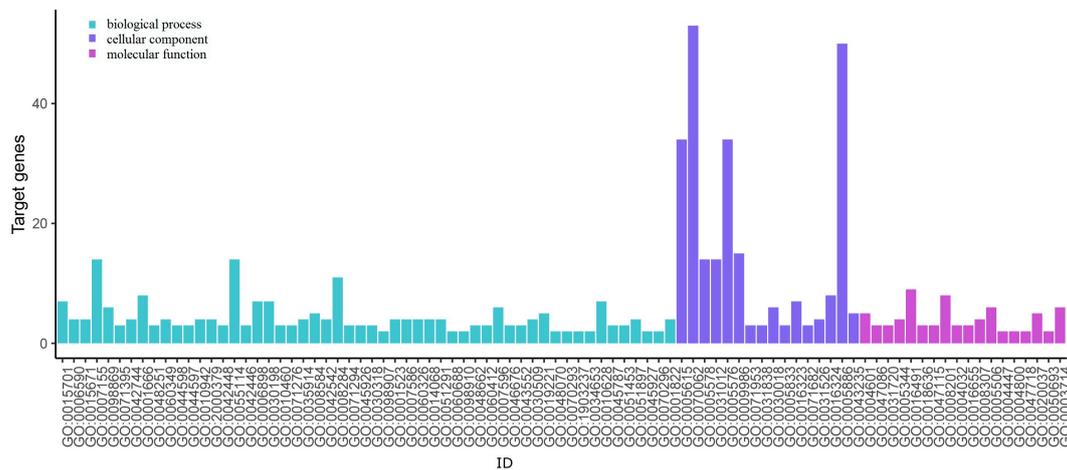
At present, most methods for identifying DEGs are focused on single gene. We propose a new idea to find the combination of genes, that is, extracting the high-density subgraphs. The high-density subgraphs are closely connected internally, and the connections with other nodes are sparse. High-density subgraphs extracted from the constructed network are shown as Figure 4. The first high-density subgraph is composed of *MUC1* [29], *LGALS3* [22], *ICAM1* [20], *TIMP1* [20], *CXCL12* [26], *MET* [22], and *FN1* [20], which are shown in Figure 4a. Consistent with what we described before, all these genes have been reported to be related to PTC. We hypothesize that *MUC1*, *LGALS3*, *ICAM1*, *TIMP1*, *CXCL12*, *MET*, and *FN1* may contribute to PTC in the form of a high-density subgraph that interact with each other. The second high-density subgraph is composed of *GABBR2* [24], *CXCR7* [52], *CCL21* [49], *CXCL12* [26], *LAPR5* [36], *NMU* [36], *GNA14* [37], *AGTR1* [34], *EDN3*, and *AVPR1A*, which are shown as Figure 4b. Among them, *GABBR2*, *CXCR7*, *CCL21*, *CXCL12*, *LAPR5*, *NMU*, *GNA14*, and *AGTR1* have been reported that they are related to PTC, and other two genes *EDN3* and *AVPR1A* have not been reported to directly lead to PTC. Therefore, the interaction of these genes in the form of a high-density subgraph may lead to PTC. The third high-density subgraph is composed of *ICAM1* [20], *EGR1* [38], *KIT* [24], *NCAM1*, *CXCL12* [26], *PROM1*, *PGF*, and *MET* [22], which are shown as Figure 4c. Among them, *ICAM1*, *EGR1*, *KIT*, *CXCL12*, and *MET* have been reported that they are related to PTC; however, *PROM1*, *PGF* and *NCAM1* have not been reported to directly lead to PTC. We hypothesize that the interaction between them in the form of high-density subgraph may lead to PTC.



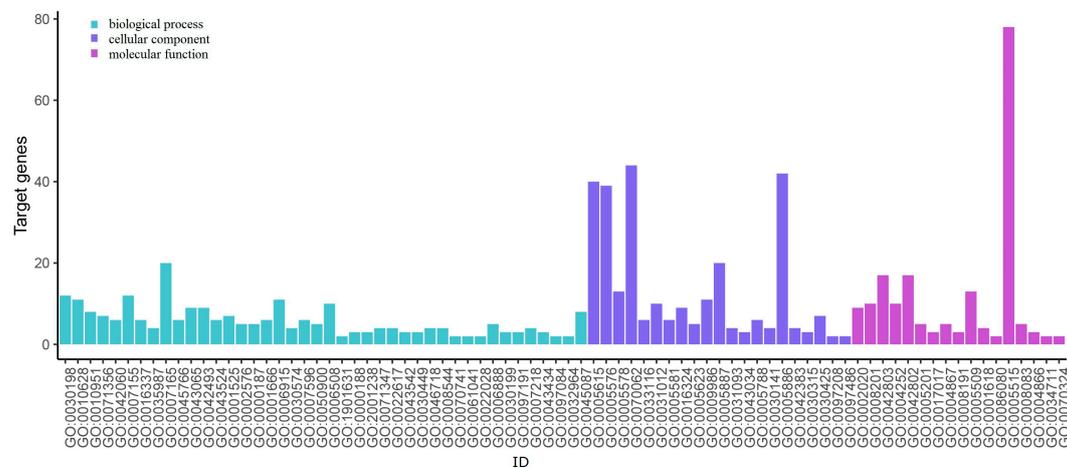
**Figure 4.** High-density subgraphs of DEGs interaction network. (a–c) show the high-density subgraphs

### 3.5. GO Term and KEGG Pathway Enrichment Analysis

The results of DEGs-related GO annotation in PTC are shown in Figures 5 and 6. The GO can be divided into three functional groups: molecular function, which indicates molecular activities of gene products; biological process, whose pathways and larger processes made up of the activities of multiple gene products; and cellular component, where gene products are active.



**Figure 5.** Up-regulated genes Gene Ontology (GO) enrichment analysis.



**Figure 6.** Down-regulated genes GO enrichment analysis.

We put the DEGs into the DAVID online tool. As is shown in Figure 5, in the biological procession, the up-regulated genes are mainly enriched in GO:0070062, GO:0005886, GO:0005615, and GO:0005576,

which correspond to extracellular exosome, plasma membrane, extracellular space, and extracellular region, respectively. As is shown in Figure 6, the down-regulated genes are mainly enriched on GO:0005515, GO:0070062, GO:0005886, GO:0005615, and GO:0005576, which correspond to protein binding, extracellular exosome, plasma membrane, extracellular space, and extracellular region, respectively. Table 3 shows the number of genes enriched on each GO term, and this table only lists the genes from each GO term that are the same as the results calculated using the new measure NCM. From the results we can see that the top 40 genes identified using the NCM measure enrich in these terms described above.

**Table 3.** Gene Ontology (GO) enrichment analysis of DEGs.

DEGs	ID	Term	Count <sup>1</sup>	Genes <sup>2</sup>
Up-regulated	GO:0070062	extracellular exosome	53	<i>TIMP1 FN1 APOE COMP SERPINA1 NMU ECM1 COL5A1 LCN2 COL1A1 PLAU MET</i>
	GO:0005886	plasma membrane	50	<i>GNA14 AGTR1 SLC4A4 IRS1 CD36 KIT LRP2 SLC26A4 ANK2</i>
	GO:0005615	extracellular space	34	<i>FMOD KIT DCN CXCL12 OGN CFD CD36</i>
	GO:0005576	extracellular region	34	<i>FMOD DCN CXCL12 OGN COL9A3 CFD MFAP4</i>
Down-regulated	GO:0005515	protein binding	78	<i>APOE COL1A1 COMP ECM1 FN1 ICAM1 LRRK2 MUC1 NMU PCSK2 PLAU SDC4 SERPINA1 TIMP1 MET SDC4 TIMP1 APOE COMP ALDH1A3 SERPINA1 FN1 MUC1 ICAM1 ECM1 COL5A1 LCN2 LRRK2 PLAU SDC4 APOE ICAM1 MET LPAR5 LRRK2 PLAU TIMP1 PCSK2 APOE COMP SERPINA1 FN1 MUC1 ICAM1 ECM1 LCN2 COL1A1 LRRK2 PLAU TIMP1 APOE COMP SERPINA1 NMU FN1 MET ECM1 COL5A1 LCN2 COL1A1 PLAU</i>
	GO:0070062	extracellular exosome	44	
	GO:0005886	plasma membrane	42	
	GO:0005615	extracellular space	40	
	GO:0005576	extracellular region	39	

<sup>1</sup> Count: the number of genes involved in the term. <sup>2</sup> Genes: genes in this term correspond to the top40 genes calculated using NCM.

By using the KOBAS online analysis database to analyze the pathway enrichment, we found that DEGs were mainly involved in 41 KEGG classical pathways, especially in cytokine-cytokine receptor interaction (12 genes), PI3K-Akt signaling pathway (12 genes) and metabolic pathways (24 genes). The three pathways with the largest number of genes connected to them are listed in Table 4. *p*-value < 0.05 is selected as the cut-off criterion to identify the enrichment pathway. These three pathways have been proven to participate in PTC. Zhang et al. [52] revealed DEGs of PTC are mainly involved in cytokine-cytokine receptor interaction pathway, which may regulate PTC growth and metastasis. Agrawal et al. [56] demonstrated that aberrant activation of the PI3K-Akt pathway plays a role in thyroid tumorigenesis. Metabolites such as neuroactive ligand-receptor interactions [57] in metabolic pathways vary widely between normal and diseased samples of PTC.

**Table 4.** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DEGs associated with PTC.

Term	ID	Input Number	<i>p</i> -Value	Input Genes
Cytokine-cytokine receptor interaction	hsa04060	12	0.0002	<i>KIT MET CXCL12 LIFR CXCL14 ACKR3 TNFRSF11B TNFRSF12A CCL18 CCL21 GHR INHBA</i>
PI3K-Akt signaling pathway	hsa04151	12	0.0017	<i>MET FN1 RELN KIT FGFR2 LPAR5 COMP COL1A1 PGF IRS1 GHR COL9A3 AKR1C3 TYMS GALNT7 TPO ADH1B LIPG ALDH1A1 GALE OGDHL TUSC30 PLA2G73 SORD CSGALNACT1 PAPS2 GATM CYP26B1 DGKI ALDH1A3 ALOX15B HGD TDO2 AOX1 HSD17B6 IMPA2</i>
Metabolic pathways	hsa01100	24	0.0267	

#### 4. Conclusions

In conclusion, a research framework was proposed to identify the important characteristic genes which may be involved in PTC progression through the integrated analysis of multiple original

microarray datasets. Based on the existing measures to quantify the topological properties of nodes, this paper proposed a new measure to comprehensively consider both local and global features of nodes. Experiments have demonstrated that this measure has important biological significance. Moreover, the availability of gene ontology and pathway enrichment analysis may provide evidence for cancer research. The results obtained in this paper may help to better understand the molecular mechanisms underlying PTC and provide a range of potential biomarkers. However, further experiments are needed to verify the results.

Our method has several advantages. First, the new measure we proposed is not simply to analyze one side of the local or global characteristics of each node, but to comprehensively analyze the topological properties of each node and the role it plays in the entire network. Second, the majority of existing studies focused on how a single characteristic gene contributes to the development of tumors in PTC, with limited research concerning the interaction of multi-genes. Herein, we analyzed the high-density subgraphs of the network constructed by DEGs to find interactions between multiple genes. Third, using network analysis methods makes the display of interactions between DEGs more intuitive and clear, and can significantly reduce the time of analysis and the difficulty of calculation.

Though our method is a beneficial exploration in detecting the DEGs, it still has several limitations, and needs further improvement and development. First, there are several parameters that need to be manually set, and no definitive standard to regulate the setting of logFC and *p*-value. Second, there are many studies to choose the DEGs at present, and no reasonable description to explain that the limma software is the best. These limitations motivate us to continue exploring in the future.

**Author Contributions:** J.S. and Q.D. jointly contributed to the design of the study. Q.D. designed and implemented the framework and new measure, performed the experiments, and drafted the manuscript. J.-X.L. and H.Z. participated in the design of the study and performed the statistical analysis. S.Y. and F.L. contributed to the data analysis. All authors approved the final manuscript.

**Funding:** This work was supported by China Postdoctoral Science Foundation (2018M642635), the National Science Foundation of China (61502272, 61572284, 61701279, 61702299), the Shandong Provincial Natural Science Foundation (ZR2017PF006, ZR2018MA019), the Project of Shandong Province Higher Educational Science and Technology Program (J18KA373, J17KA063).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fan, M.; Li, X.; Jiang, W.; Huang, Y.; Li, J.; Wang, Z. A long non-coding RNA, PTCSC3, as a tumor suppressor and a target of miRNAs in thyroid cancer cells. *Exp. Ther. Med.* **2013**, *5*, 1143–1146. [[CrossRef](#)] [[PubMed](#)]
2. Kim, W.Y.; Lee, J.B.; Jung, S.P.; Kim, H.Y.; Woo, S.U.; Son, G.S.; Bae, J.W. Gene expression profiles of papillary thyroid microcarcinoma. *Int. Surg.* **2017**, *102*, 39–46. [[CrossRef](#)]
3. El Naggar, A.K. Gene expression profiling of differentiated thyroid neoplasms: Diagnostic and clinical implications. *Clin. Cancer Res.* **2004**, *10*, 6586–6597.
4. Barbara, J.; Malgorzata, W.; Krzysztof, F.; Krzysztof, S.; Michal, J.; Malgorzata, O.-W.; Jan, W.; Agnieszka, C.; Ewa, C.; Dariusz, L. Gene expression profile of papillary thyroid cancer: Sources of variability and diagnostic implications. *Cancer Res.* **2005**, *65*, 1587–1597.
5. He, Z.; Sun, M.; Ke, Y.; Lin, R.; Xiao, Y.; Zhou, S.; Zhao, H.; Wang, Y.; Zhou, F.; Zhou, Y. Identifying biomarkers of papillary renal cell carcinoma associated with pathological stage by weighted gene co-expression network analysis. *Oncotarget* **2017**, *8*, 27904–27914. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, J.; Li, J.; Deng, H.W. Identifying gene interaction enrichment for gene expression data. *PLoS ONE* **2009**, *4*, e8064. [[CrossRef](#)]
7. Rizzolo, K.; Kumar, A.; Kakihara, Y.; Phanse, S.; Minic, Z.; Snider, J.; Stagljar, I.; Zilles, S.; Babu, M.; Houry, W.A. Systems analysis of the genetic interaction network of yeast molecular chaperones. *Mol. Omics* **2018**, *14*, 82–94. [[CrossRef](#)]
8. Wei, P.J.; Zhang, D.; Xia, J.; Zheng, C.H. LNDriver: Identifying driver genes by integrating mutation and expression data based on gene-gene interaction network. *BMC Bioinform.* **2016**, *17*, 467. [[CrossRef](#)]

9. He, H.; Jazdzewski, K.; Li, W.; Liyanarachchi, S.; Nagy, R.; Volinia, S.; Calin, G.A.; Liu, C.G.; Franssila, K.; Suster, S. The role of microRNA genes in papillary thyroid carcinoma. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 19075–19080. [[CrossRef](#)]
10. He, W.; Qi, B.; Zhou, Q.; Lu, C.; Huang, Q.; Xian, L.; Chen, M. Key genes and pathways in thyroid cancer based on gene set enrichment analysis. *Oncol. Rep.* **2013**, *30*, 1391–1397. [[CrossRef](#)]
11. Tomás, G.; Tarabichi, M.; Gacquer, D.; Hébrant, A.; Dom, G.; Dumont, J.E.; Keutgen, X.; Rd, F.T.; Maenhaut, C.; Detours, V. A general method to derive robust organ-specific gene expression-based differentiation indices: Application to thyroid cancer diagnostic. *Oncogene* **2012**, *31*, 4490–4498. [[CrossRef](#)] [[PubMed](#)]
12. Rusinek, D.; Swierniak, M.; Chmielik, E.; Kowal, M.; Kowalska, M.; Cyplinska, R.; Czarniecka, A.K.; Piglowski, W.; Korfanty, J.; Chekan, M. BRAFV600E-associated gene expression profile: Early changes in the transcriptome, based on a transgenic mouse model of papillary thyroid carcinoma. *PLoS ONE* **2015**, *10*, e0143688. [[CrossRef](#)] [[PubMed](#)]
13. Smyth, G.K.; Michaud, J.; Scott, H.S. *Use of within-Array Replicate Spots for Assessing Differential Expression in Microarray Experiments*; Oxford University Press: Oxford, UK, 2005.
14. Chen, W.; Liu, Q.; Lv, Y.; Xu, D.; Chen, W.; Yu, J. Special role of *JUN* in papillary thyroid carcinoma based on bioinformatics analysis. *World J. Surg. Oncol.* **2017**, *15*, 119. [[CrossRef](#)] [[PubMed](#)]
15. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P. The string database in 2017: Quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **2017**, *45*, D362–D368. [[CrossRef](#)] [[PubMed](#)]
16. Chin, C.H.; Chen, S.H.; Wu, H.H.; Ho, C.W.; Ko, M.T.; Lin, C.Y. CytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **2014**, *8* (Suppl. 4), S11. [[CrossRef](#)]
17. Tadaka, S.; Kinoshita, K. NCMine: Core-peripheral based functional module detection using near-clique mining. *Bioinformatics* **2016**, *32*, 3454–3460. [[CrossRef](#)]
18. Dennis, G.; Sherman, B.T.; Hosack, D.A.; Yang, J.; Gao, W.; Lane, H.C.; Lempicki, R.A. DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **2003**, *4*, P3. [[CrossRef](#)]
19. Kanehisa, M.; Furumichi, M.; Mao, T.; Sato, Y.; Morishima, K. Kegg: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [[CrossRef](#)]
20. Qiu, J.; Zhang, W.; Zang, C.; Liu, X.; Liu, F.; Ge, R.; Sun, Y.; Xia, Q. Identification of key genes and miRNAs markers of papillary thyroid cancer. *Biol. Res.* **2018**, *51*, 45. [[CrossRef](#)]
21. Qiu, J.; Zhang, W.; Xia, Q.; Liu, F.; Li, L.; Zhao, S.; Gao, X.; Zang, C.; Ge, R.; Sun, Y. RNA sequencing identifies crucial genes in papillary thyroid carcinoma (PTC) progression. *Exp. Mol. Pathol.* **2016**, *100*, 151–159. [[CrossRef](#)]
22. Qu, T.; Li, Y.P.; Li, X.H.; Chen, Y. Identification of potential biomarkers and drugs for papillary thyroid cancer based on gene expression profile analysis. *Mol. Med. Rep.* **2016**, *14*. [[CrossRef](#)]
23. Rosignolo, F.; Sponziello, M.; Durante, C.; Puppini, C.; Mio, C.; Baldan, F.; Di, L.C.; Russo, D.; Filetti, S.; Damante, G. Expression of *PAX8* target genes in papillary thyroid carcinoma. *PLoS ONE* **2016**, *11*, e0156658. [[CrossRef](#)] [[PubMed](#)]
24. Schulten, H.J.; Al-Mansouri, Z.; Baghallab, I.; Bagatian, N.; Subhi, O.; Karim, S.; Al-Aradati, H.; Al-Mutawa, A.; Johary, A.; Meccawy, A.A. Comparison of microarray expression profiles between follicular variant of papillary thyroid carcinomas and follicular adenomas of the thyroid. *BMC Genom.* **2015**, *16*, S7. [[CrossRef](#)] [[PubMed](#)]
25. Hyun Sook, K.; Do Hyung, K.; Yeon, K.J.; Nam Ho, J.; In Kyu, L.; Gu, B.J.; Eui Dal, J. Microarray analysis of papillary thyroid cancers in Korean. *Korean J. Intern. Med.* **2010**, *25*, 399–407.
26. Liang, W.; Sun, F. Identification of key genes of papillary thyroid cancer using integrated bioinformatics analysis. *J. Endocrinol. Investig.* **2018**, *41*, 1237–1245. [[CrossRef](#)] [[PubMed](#)]
27. Zhang, K.; Liu, J.; Li, C.; Peng, X.; Li, H.; Li, Z. Identification and validation of potential target genes in papillary thyroid cancer. *Eur. J. Pharmacol.* **2019**, *843*, 217–225. [[CrossRef](#)] [[PubMed](#)]
28. Weber, F.; Shen, L.; Aldred, M.A.; Morrison, C.D.; Frilling, A.; Saji, M.; Schuppert, F.; Broelsch, C.E.; Ringel, M.D.; Eng, C. Genetic classification of benign and malignant thyroid follicular neoplasia based on a three-gene combination. *J. Clin. Endocrinol. Metab.* **2005**, *90*, 2512–2521. [[CrossRef](#)] [[PubMed](#)]
29. Li, Q.; Jin, W.X.; Jin, Y.X.; Zheng, Z.C.; Zhou, X.F.; Wang, Q.X.; Ye, D.R.; Sun, Y.H.; Zhang, X.H.; Wang, O.C. Clinical effect of *MUC1* and its relevance to *BRAF* V600E mutation in papillary thyroid carcinoma: A case–control study. *Cancer Manag. Res.* **2018**, *10*, 1351–1358. [[CrossRef](#)]

30. Wei, X.; Zhang, Y.; Yu, S.; Li, S.; Jiang, W.; Zhu, Y.; Xu, Y.; Yang, C.; Tian, G.; Mi, J. PDLIM5 identified by label-free quantitative proteomics as a potential novel biomarker of papillary thyroid carcinoma. *Biochem. Biophys. Res. Commun.* **2018**, *499*, 338–344. [[CrossRef](#)]
31. Katsuhiko, T.; Hiroshi, S.; Junichi, K.; Tsunehisa, N.; Sumiko, O.; Yutaka, Y.; Shigeru, Y. Inhibition of infiltration and angiogenesis by thrombospondin-1 in papillary thyroid carcinoma. *Clin. Cancer Res.* **2002**, *8*, 1125–1131.
32. Arnaldi, L.A.; Borra RCMacieli, R.M.; Cerutti, J.M. Gene expression profiles reveal that *DCN*, *DIO1*, and *DIO2* are underexpressed in benign and malignant thyroid tumors. *Thyroid* **2005**, *15*, 210–221. [[CrossRef](#)]
33. Lisi, S.; Botta, R.; Rotondo Dottore, G.; Leo, M.; Latrofa, F.; Vitti, P.; Marinò, M. Intracellular retention of thyroglobulin in the absence of the low-density lipoprotein receptor-associated protein (RAP) is likely due to premature binding to megalin in the biosynthetic pathway. *J. Endocrinol. Investig.* **2016**, *39*, 1039–1044. [[CrossRef](#)] [[PubMed](#)]
34. Guo, S.; Yan, F.; Xu, J.; Bao, Y.; Zhu, J.; Wang, X.; Wu, J.; Li, Y.; Pu, W.; Liu, Y.; et al. Identification and validation of the methylation biomarkers of non-small cell lung cancer (NSCLC). *Clin. Epigenet.* **2015**, *7*, 3. [[CrossRef](#)] [[PubMed](#)]
35. Tan, J.; Qian, X.; Song, B.; An, X.; Cai, T.; Zuo, Z.; Ding, D.; Lu, Y.; Li, H. Integrated bioinformatics analysis reveals that the expression of cathepsin S is associated with lymph node metastasis and poor prognosis in papillary thyroid cancer. *Oncol. Rep.* **2018**, *40*, 111–122. [[PubMed](#)]
36. Tang, J.; Kong, D.; Cui, Q.; Wang, K.; Zhang, D.; Yuan, Q.; Liao, X.; Gong, Y.; Wu, G. Bioinformatic analysis and identification of potential prognostic microRNAs and mRNAs in thyroid cancer. *PeerJ* **2018**, *6*, e4674. [[CrossRef](#)] [[PubMed](#)]
37. Kai-Chun, C.; Chun-Tzu, H.; Kai-Yuan, C.; Kuo-Jen, C.; Wen-Chuan, W.; Jau-Ling, S.; Yu-Jen, W.; Cheng-Hsien, C. Proteomic surveillance of putative new autoantigens in thyroid orbitopathy. *Br. J. Ophthalmol.* **2015**, *99*, 1571–1576.
38. Lim, A.S.P.; Gaiteri, C.; Yu, L.; Sohail, S.; Swardfager, W.; Tasaki, S.; Schneider, J.A.; Paquet, C.; Stuss, D.T.; Masellis, M.; et al. Seasonal plasticity of cognition and related biological measures in adults with and without Alzheimer disease: Analysis of multiple cohorts. *PLoS Med.* **2018**, *15*, e1002647. [[CrossRef](#)] [[PubMed](#)]
39. Stein, L. Copy number and gene expression alterations in radiation-induced papillary thyroid carcinoma from Chernobyl pediatric patients. *Thyroid* **2010**, *20*, 475–487. [[CrossRef](#)]
40. Maro, G.D.; Salerno, P.; Unger, K.; Orlandella, F.M.; Monaco, M.; Chiappetta, G.; Thomas, G.; Oczko-Wojciechowska, M.; Masullo, M.; Jarzab, B. Anterior gradient protein 2 promotes survival, migration and invasion of papillary thyroid carcinoma cells. *Mol. Cancer* **2014**, *13*, 160. [[CrossRef](#)]
41. Griffith, O.L.; Adrienne, M.; Jones, S.J.M.; Wiseman, S.M. Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. *J. Clin. Oncol.* **2006**, *24*, 5043–5051. [[CrossRef](#)]
42. Tailor, P.D.; Kodeboyina, S.K.; Bai, S.; Patel, N.; Sharma, S.; Ratnani, A.; Copland, J.A.; She, J.X.; Sharma, A. Diagnostic and prognostic biomarker potential of kallikrein family genes in different cancer types. *Oncotarget* **2018**, *9*, 17876–17888. [[CrossRef](#)] [[PubMed](#)]
43. Ito, Y.; Yoshida, H.; Tomoda, C.; Uruno, T.; Miya, A.; Kobayashi, K.; Matsuzuka, F.; Kakudo, K.; Kuma, K.; Miyauchi, A. Expression of S100A2 and S100A6 in thyroid carcinomas. *Histopathology* **2010**, *46*, 569–575. [[CrossRef](#)] [[PubMed](#)]
44. Nasr, M.R.; Mukhopadhyay, S.; Zhang, S.; Katzenstein, A.A. Immunohistochemical markers in diagnosis of papillary thyroid carcinoma: Immunohistochemical markers in diagnosis of papillary thyroid carcinoma: Utility of HBME1 combined with CK19 immunostaining. *Mod. Pathol.* **2006**, *19*, 1631–1637. [[CrossRef](#)]
45. Electron, K.; Miao, P.; Emily, R.; Quan-Yang, D.; Clark, O.H.; Alex, M.M. ECM1 and TMPRSS4 are diagnostic markers of malignant thyroid neoplasms and improve the accuracy of fine needle aspiration biopsy. *Ann. Surg.* **2005**, *242*, 353–363.
46. Abend, M.; Pfeiffer, R.M.; Ruf, C.; Hatch, M.; Bogdanova, T.I.; Tronko, M.D.; Hartmann, J.; Meineke, V.; Mabuchi, K.; Brenner, A.V. Iodine-131 dose-dependent gene expression: Alterations in both normal and tumour thyroid tissues of post-Chernobyl thyroid cancers. *Br. J. Cancer* **2013**, *109*, 2286–2294. [[CrossRef](#)] [[PubMed](#)]
47. Weinberger, P.; Ponny, S.R.; Xu, H.; Bai, S.; Smallridge, R.; Copland, J.; Sharma, A. Cell cycle M-phase genes are highly upregulated in anaplastic thyroid carcinoma. *Thyroid* **2017**, *27*, 236–252. [[CrossRef](#)] [[PubMed](#)]

48. Wu, C.C.; Lin, J.D.; Chen, J.T.; Chang, C.M.; Weng, H.F.; Hsueh, C.; Chien, H.P.; Yu, J.S. Integrated analysis of fine-needle-aspiration cystic fluid proteome, cancer cell secretome, and public transcriptome datasets for papillary thyroid cancer biomarker discovery. *Oncotarget* **2018**, *9*, 12079–12100. [[CrossRef](#)] [[PubMed](#)]
49. Smallridge, R.C.; Ana-Maria, C.; Yan, W.; Asmann, Y.W.; Casler, J.D.; Serie, D.J.; Reddi, H.V.; Cradic, K.W.; Michael, R.; Grebe, S.K.; et al. RNA sequencing identifies multiple fusion transcripts, differentially expressed genes, and reduced expression of immune function genes in *BRAF* (V600E) mutant vs *BRAF* wild-type papillary thyroid carcinoma. *J. Clin. Endocrinol. Metab.* **2014**, *99*, E338–E347. [[CrossRef](#)]
50. Mateus Camargo, B.F.; Fabio Albuquerque, M.; Clóvis Ant?Nio, P.; Silvia Regina, R.; Luiz Paulo, K. High diagnostic accuracy based on *CLDN10*, *HMGA2*, and *LAMB3* transcripts in papillary thyroid carcinoma. *J. Clin. Endocrinol. Metab.* **2015**, *100*, 890–899.
51. Jiao, F.; Lv, H.; Guan, H.; Ma, X.; Ji, M.; He, N.; Shi, B.; Peng, H. High diagnostic accuracy based on *CLDN10*, *HMGA2*, and *LAMB3* transcripts in papillary thyroid carcinoma. *J. Clin. Endocrinol. Metab. BMC Cancer* **2013**, *13*, 462.
52. Zhang, H.; Teng, X.; Liu, Z.; Lei, Z.; Zhen, L. Gene expression profile analyze the molecular mechanism of CXCR7 regulating papillary thyroid carcinoma growth and metastasis. *J. Exp. Clin. Cancer Res.* **2015**, *34*, 1–12. [[CrossRef](#)] [[PubMed](#)]
53. Janssen, S.F.; Gorgels, T.G.; Spek, P.J.V.D.; Jansonius, N.M.; Bergen, A.A. In silico analysis of the molecular machinery underlying aqueous humor production: Potential implications for glaucoma. *J. Clin. Bioinform.* **2013**, *3*, 21. [[CrossRef](#)] [[PubMed](#)]
54. Dziegiel, P.; Pula, B.; Kobierzycki, C.; Stasiolek, M.; Podhorska-Okolow, M. *The Role of Metallothioneins in Carcinogenesis*; Springer International Publishing: Berlin, Germany, 2016; Volume 218, pp. 29–63.
55. Passon, N.; Bregant, E.; Sponziello, M.; Dima, M.; Rosignolo, F.; Durante, C.; Celano, M.; Russo, D.; Filetti, S.; Damante, G. Somatic amplifications and deletions in genome of papillary thyroid carcinomas. *Endocrine* **2015**, *50*, 1–12. [[CrossRef](#)] [[PubMed](#)]
56. Agrawal, N.; Akbani, R.; Aksoy, B.A.; Ally, A.; Arachchi, H.; Asa, S.; Auman, J.T.; Balasundaram, M.; Balu, S.; Baylin, S. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **2014**, *159*, 676–690. [[CrossRef](#)] [[PubMed](#)]
57. Shang, X.; Zhong, X.; Tian, X. Metabolomics of papillary thyroid carcinoma tissues: Potential biomarkers for diagnosis and promising targets for therapy. *Tumor Biol.* **2016**, *37*, 11163–11175. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).