

# Network medicine approaches for identification of novel prognostic systems biomarkers and drug candidates for papillary thyroid carcinoma

Medi Kori<sup>1,2</sup> | Kubra Temiz<sup>3</sup> | Esra Gov<sup>3</sup> 

<sup>1</sup>Faculty of Health Sciences, Acibadem Mehmet Ali Aydınlar University, İstanbul, Türkiye

<sup>2</sup>Department of Bioengineering, Marmara University, İstanbul, Türkiye

<sup>3</sup>Department of Bioengineering, Adana Alparslan Türkeş Science and Technology University, Adana, Türkiye

## Correspondence

Esra Gov, Department of Bioengineering, Faculty of Engineering, Adana Alparslan Türkeş Science and Technology University, Building M1, Office: 202, Sarıcam, 01250 Adana, Türkiye.

Email: [medi.kori@acibadem.edu.tr](mailto:medi.kori@acibadem.edu.tr)

## Funding information

Türkiye Sağlık Enstitüleri Başkanlığı, Grant/Award Number: 2019-TA-01/3436

## Abstract

Papillary thyroid carcinoma (PTC) is one of the most common endocrine carcinomas worldwide and the aetiology of this cancer is still not well understood. Therefore, it remains important to understand the disease mechanism and find prognostic biomarkers and/or drug candidates for PTC. Compared with approaches based on single-gene assessment, network medicine analysis offers great promise to address this need. Accordingly, in the present study, we performed differential co-expressed network analysis using five transcriptome datasets in patients with PTC and healthy controls. Following meta-analysis of the transcriptome datasets, we uncovered common differentially expressed genes (DEGs) for PTC and, using these genes as proxies, found a highly clustered differentially expressed co-expressed module: a 'PTC-module'. Using independent data, we demonstrated the high prognostic capacity of the PTC-module and designated this module as a prognostic systems biomarker. In addition, using the nodes of the PTC-module, we performed drug repurposing and text mining analyzes to identify novel drug candidates for the disease. We performed molecular docking simulations, and identified: 4-demethoxydaunorubicin hydrochloride, AS605240, BRD-A60245366, ER 27319 maleate, sinensetin, and TWS119 as novel drug candidates whose efficacy was also confirmed by in silico analyzes. Consequently, we have highlighted here the need for differential co-expression analysis to gain a systems-level understanding of a complex disease, and we provide candidate prognostic systems biomarker and novel drugs for PTC.

## KEYWORDS

co-expressed module, differential co-expression network analysis, drug repurposing, papillary thyroid carcinoma, prognostic systems biomarkers, text mining

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Recent cancer statistics show that thyroid cancer is the most common endocrine cancer encountered by mankind. In 2022, it was estimated that 43,800 patients were diagnosed with thyroid cancer, and thyroid cancer was responsible for 2230 deaths in the United States.<sup>1</sup> Papillary, follicular, medullary, and undifferentiated or anaplastic thyroid carcinomas are the major histologic groups of thyroid tumours. Among the various subgroups of thyroid cancer, papillary thyroid carcinoma (PTC) accounts for the largest proportion of all thyroid carcinomas. Notably, the prevalence of PTC is approximately 85% when all thyroid carcinomas are considered.<sup>2</sup> The aetiology of thyroid cancer is not well understood. Therefore, further studies are needed to uncover the underlying mechanisms of the disease, correct over-diagnosis, and find prognostic and/or drug candidates to reduce the prevalence of PTC.

Logarithmic advances in high-throughput sequencing and screening methods in recent decades have produced a substantial amount of X-ome data at various molecular levels, enabling researchers to perform various bioinformatics approaches for different diseases<sup>3,4</sup> and also for PTC.<sup>5</sup> However, some of these studies take only limited account of the reality of the molecular biochemistry of the organism. To discover specific disease biomarkers or drug targets, it is essential to evaluate the entire physical and functional architecture of the organism, because the development of an abnormal phenotype (i.e. disease) is not the result of a single gene, but rather the result of complex gene interactions. One scientific field, network medicine, allows researchers to uncover these complex interactions between biomolecules of a given phenotype in a holistic view.<sup>6</sup>

Gene and protein networks provide valuable data for molecular interactions within the organism, and co-expression networks represent significantly co-regulated groups of genes (i.e. modules). Any differentiation in gene correlations between different phenotypes can provide clues to the phenotype and supports the discovery of systems biomarkers.<sup>7</sup> To reveal co-expression relationships among genes in thyroid cancer, weighted gene co-expression network analysis (WGCNA) has been performed in a limited number of studies.<sup>8,9</sup> WGCNA identifies gene modules by hierarchical clustering. Due to drawbacks of the algorithm, there is still no major standard for WGCNA.<sup>10</sup> Therefore, in this study, instead of using the WGCNA pipeline, we used a different bioinformatics pipeline whose efficiency has also been demonstrated in various cancers, including ovarian cancer,<sup>11</sup> cervical cancer,<sup>12</sup> ovarian cancer stem cells,<sup>13</sup> and gastric cancer.<sup>14</sup> To increase the statistical power of our analysis, unlike other studies in this study, we performed a meta-analysis based on the integration of information from multiple studies.

Accordingly, in the present study, five transcriptome datasets from different cohorts were statistically meta-analysed to identify differentially expressed genes (DEGs) in PTC. By using culminated common DEGs between datasets, we identified co-expressed gene

pairs that differ between PTC and control phenotypes. After identifying the major significantly co-expressed gene pairs, we constructed co-expression networks and determined co-expressed modules. We evaluated the prognostic capacity of the culminated modules, and the nodes of the modules that have high prognostic capacity were used as proxies to find drug candidates. Text mining and molecular docking analyzes were used to demonstrate the novelty and efficacy of the drug candidates.

## 2 | MATERIALS AND METHODS

### 2.1 | Selection of transcriptome datasets

Five gene expression datasets from PTC, including GSE60542,<sup>15</sup> GSE3678, GSE29265, GSE35570<sup>16</sup> and GSE33630,<sup>17</sup> were obtained from the Gene Expression Omnibus (GEO)<sup>18</sup> database. When selecting the datasets, we paid attention to the following three parameters: (i) the selected datasets were from the same platform (i.e. Affymetrix array), (ii) each phenotype contained at least three samples and (iii) the cancer samples belonged to papillary histology. The five datasets obtained included 141 PTC patients and 153 matched non-tumour controls and were used to identify co-expressed modules during the study. In addition, the clinical information and expression values of an independent thyroid cancer dataset (THCA) were extracted from The Cancer Genome Atlas (TCGA)<sup>19</sup> to pre-clinically validate the prognostic performance of the obtained modules. Since we focus on PTC, we extracted only papillary cancer cases from the THCA-TCGA, representing a total number of 494 PTC samples.

### 2.2 | Identification of DEGs and pathway over-representation analyses

In this study, we used a well-established statistical analysis procedure<sup>20,21</sup> to identify DEGs. Briefly, raw data from each of the five datasets were normalized by calculating the Robust Multi-Array Average (RMA) expression measure<sup>22</sup> implemented in the Affy package<sup>23</sup> of the R/Bioconductor platform (version 4.0.2).<sup>24</sup> DEGs were identified from normalized expression values using the Linear Models for Microarray Data package (LIMMA).<sup>25</sup> The Benjamini-Hochberg method was used to control for false discovery rate (FDR). The adjusted *p* value <0.05 was used as a cut-off value to determine the statistical significance of the DEGs. To determine the regulatory patterns of DEGs (i.e. up- and down-regulation), the fold-change threshold was used, with a two-fold change accepted as statistically significant. Each data set was analysed independently, and results were comparatively analysed to identify common signatures from these independent studies.

Gene set overrepresentation analyses were performed using the functional annotation tool ConsensusPathDB<sup>26</sup> to reveal associated

biological pathways of common DEGs. The analyses preferentially used the Kyoto Encyclopedia of Genes and Genomes (KEGG)<sup>27</sup> and Reactome<sup>28</sup> as data sources for the pathways. The *p*-values were determined using Fisher's Exact Test, and FDR was applied to control the *p*-values. An adjusted *p*-value < 0.05 was considered statistically significant in overrepresentation analyses.

### 2.3 | Identification of co-expressed genes in cancer and control samples

To identify co-expressed genes between PTC and control samples, an established method<sup>11</sup> was performed using culminated common DEGs. According to this method, *z*-score normalization was performed to eliminate batch effects of each data set, and if the common DEG occurred more than once in the same dataset, the mean expression values were calculated. Then, the correlation patterns of each common gene pair in each condition (i.e. PTC and control) were calculated using correlation coefficients. A cut-off value for the correlation coefficient was set, corresponding to a *p* value of <0.05 and was used to determine the statistical significance of the pairwise correlations.

### 2.4 | Construction of co-expression networks and determination of differential co-expressed network modules

To evaluate the major significantly co-expressed genes, *p*-critical values for common DEGs were calculated by integrating the obtained pairwise gene expression correlation values with Spearman correlation coefficients (SCCs). An equation for determining the *p*-critical values (*P*<sub>critic</sub>), which is the cut-off value of the correlation coefficient, can be found in Equation 1 (Equation 1).

$$p_{critic} = \text{mean SCC} + 1.96 * \text{std of SCC} \quad (1)$$

The classification of genes and the significance of expression levels that differ in tumour and non-tumour samples was determined by the defined threshold condition (parameter  $\epsilon$ ). This threshold helps to identify differential co-expression by revealing gene pairs that are significantly linked in the tumour samples but have an opposite direction of correlation in one of the cases. For an equation and consideration of the  $\epsilon$  parameter in this study, we used Equation 2 (Equation 2). SCC<sub>d</sub> represents the correlation coefficient of gene pairs in tumour tissue and SCC<sub>h</sub> represents the correlation coefficient of gene pairs in healthy tissue.

$$\epsilon = |\text{SCC}_d - \text{SCC}_h| > 0.5 \quad (2)$$

### 2.5 | Prognostic performance analysis of differential co-expressed modules

The prognostic abilities of significant co-expressed modules were determined using independent RNA-Seq data from TGCA

(THCA-TGCA).<sup>19</sup> Clinical information on the samples was collected by TCGA to evaluate the prognostic performance of the modules. Because the most important prognostic factor in all cancers is the stage of the patient, we considered the stage of the patient in the analysis of prognostic performance. Of the 494 PTC samples we obtained from TCGA, 278 samples were stage I, 51 samples were stage II, 110 samples were stage III, 53 samples were stage IV (including IVA and IVC), and the stage was unknown for two samples, so these two samples were not included in the analyses of prognostic performance. The prognostic abilities of the modules were assessed using Kaplan–Meier plots and the log-rank test. All analyses were performed using the Survival package in R/Bioconductor (version 4.0.2).<sup>24</sup> Differential co-expressed modules with a log-rank *p*-value of <0.05 were considered statistically significant and accepted as prognostic co-expressed modules.

### 2.6 | Drug repurposing and text mining analysis

Drug repurposing analysis was performed via L1000CDS2: the LINCS L1000 search engine for characteristic directional signatures<sup>29</sup> using genes within the prognostic gene module. The analysis considered drugs that had an inverse effect on an entered gene expression signature, and the small molecules/drugs with the highest score (i.e. top 50) were selected for further investigation (i.e. text mining and molecular docking analysis).

To determine whether the identified drug candidates had been previously associated with PTC, the identified drug candidates were reviewed in the literature with Python using text mining techniques. Accordingly, the identified drug candidates were searched separately using the following four keywords 'drug candidate', 'drug candidate + thyroid cancer', 'drug candidate + thyroid carcinoma', and 'drug candidate + cancer' in the article abstracts and the results were analysed using the BioPython package.<sup>30</sup> We then, calculated the term frequency (TF) and inverse document frequency (IDF). The TF-IDF is calculated by multiplying the TF and IDF scores. Drugs with a '0' TF-IDF score were considered as novel drug candidates in the study.

### 2.7 | Molecular docking simulations

The three-dimensional (3D) crystal structure of the target proteins of the drug candidates was taken from the Protein Data Bank (PDB),<sup>31</sup> while the structures of the drug candidates were taken from the PubChem database<sup>32</sup> when available. Molecular docking analyzes were performed using AutoDock Vina software.<sup>33</sup> For docking simulations, we used previously known binding residues of the target proteins and calculated binding affinities (kcal/mol) to determine binding significance. In addition, the drugs with a higher TF-IDF value were considered as positive controls. Molecular docking simulations were also performed for these drugs to compare the binding significance of the novel drug candidates.

### 3 | RESULTS

#### 3.1 | Transcriptomic signatures of papillary thyroid cancer: DEGs and their biological insights

Five microarray gene expression datasets were analysed statistically and individually to identify DEGs. The number of DEGs in each data set showed a wide range from 955 to 1907 genes, and the highest number of DEGs was identified in GSE35570. While the number of up-regulated genes in datasets GSE33630 and GSE35570 was lower than the number of down-regulated genes, data sets GSE29265, GSE3678 and GSE60542 showed a profile with a higher number of down-regulated genes (Figure 1A). Comparative analysis of the resulting DEGs showed that 534 DEGs were common in five transcriptome datasets. To ensure consistency and robustness of the analysis, further analysis was performed using these common DEGs (Figure 1B).

Overrepresentation analysis showed that common DEGs were associated with multiple cancer-associated pathways (i.e. TGF-beta pathway, PI3K-Akt pathway and focal adhesion) and small cell lung cancer. Interestingly, we found that many collagen-associated pathways, including collagen formation, collagen chain trimerization, collagen biosynthesis and modifying enzymes, collagen fibril assembly and other multimeric structures were significantly associated with PTC. In addition, signal transduction, extracellular matrix organization, axon guidance, and nervous system development were found to be remarkably related to the disease (Figure 1C).

#### 3.2 | Differential co-expressed modules for papillary thyroid cancer

After identifying common DEGs in five databases, we used SCCs to identify key gene pairs that are co-expressed and constructed networks around them. Within the constructed co-expressed networks, a total of 12 co-expressed modules were identified. However, considering the parameters we set for detecting highly clustered co-expressed modules, only one module named 'PTC-module' was identified within the set criteria (Figure 2A).

The PTC-module consists of 21 nodes with 45 edges and has a network density of 69%. The co-expression patterns of genes between the control and disease states showed that in the control state, 14 (9.7%) gene patterns were positively correlated and eight (5%) were negatively correlated, whereas in the disease state, 69 (48%) gene patterns were positively correlated, and 73 (51%) gene patterns were negatively correlated, indicating that disease development affects the expression of gene patterns by 83% (Figure 2B).

#### 3.3 | Prognostic performance of the co-expressed module

To pre-clinically evaluate the prognostic ability of the PTC-module, an independent RNA-Seq dataset (THCA-TCGA) was used, and the

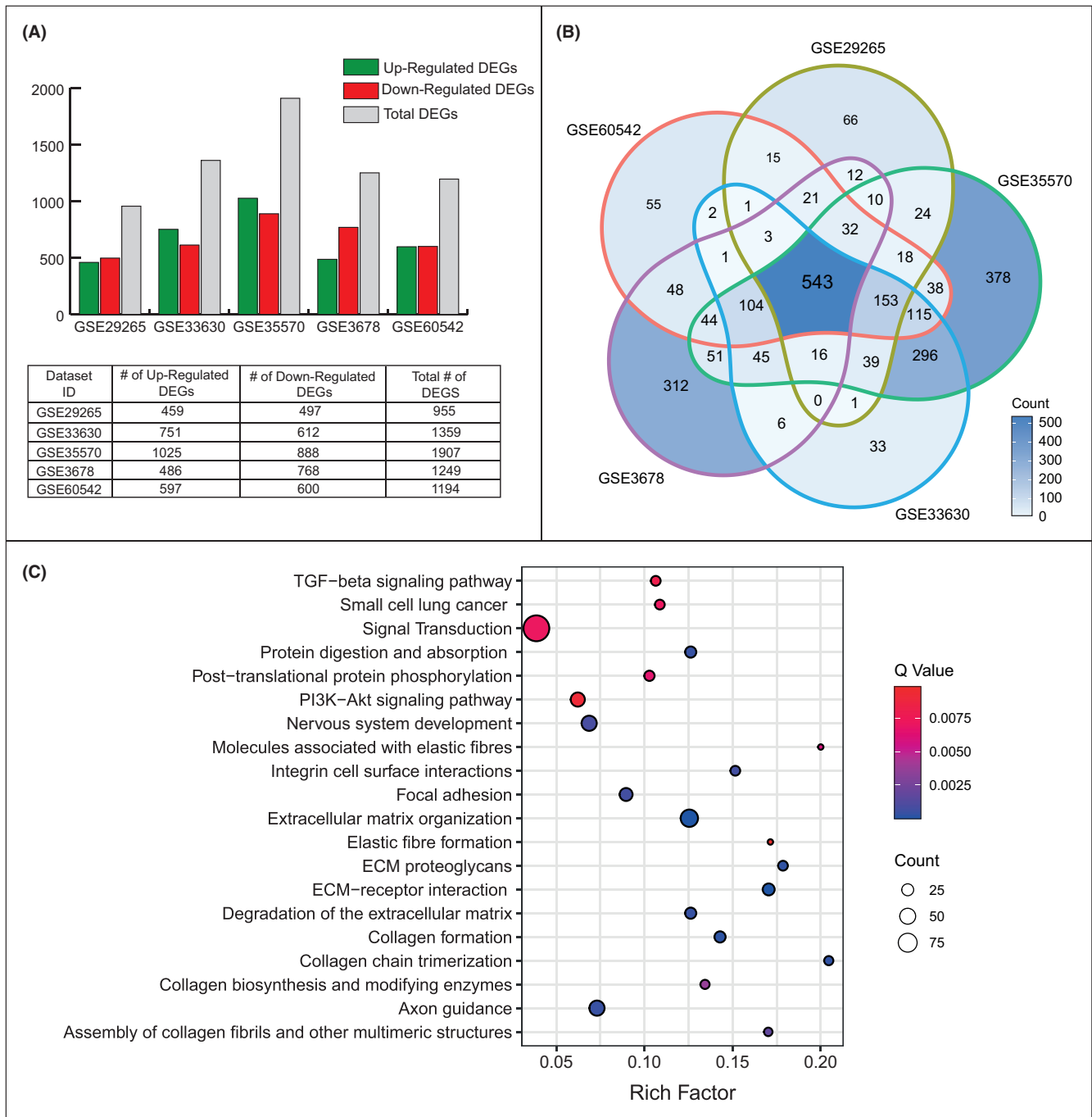
prognostic ability of the module was evaluated based on the disease stage of the patients using the Kaplan–Meier plot and the log-rank test. Since we know that the most important prognostic factor in cancer patients is disease stage, we included this information in the analysis of prognostic performance. The analyses showed that the prognostic power of the PTC-module (log-rank  $p$ -value =  $1.90 \times 10^{-4}$ ) was statistically significant, so we found that this module had a high impact on the overall survival of patients according to cancer stage (Figure 3) and was considered as a prognostic systems biomarker in the study.

#### 3.4 | Identification of novel candidate repurposed therapeutic agents

We implemented drug repurposing approach using genes in the PTC-module using the L100CDS<sup>2</sup> search engine. Subsequent analysis revealed that, 43 drugs had significant potential to reverse the expression profiles of differentially co-expressed genes. To reveal the novelty of these drug candidates, the discovered drugs were examined by text-mining analysis (Table S1). Doxorubicin hydrochloride, doxorubicin and dorsomorphin dihydrochloride yielded the higher TF-IDF values than the other drug candidates. The drug candidates that have not yet been studied in thyroid cancer (i.e. the TF-IDF values are '0') were accepted as novel drug candidates for PTC (Figure 4A).

#### 3.5 | Molecular docking simulations

We performed molecular docking simulations to evaluate the binding affinities of the novel drugs. Of the 12 novel drugs, no 3D-structure could be found for BD 1063 dihydrochloride, so it was not included in the docking analysis. We examined the 3D structures of PTC-module proteins targeting 11 novel drug candidates and found suitable structures for the following proteins: (i) CD55 (PDB structure = 1H04),<sup>34</sup> (ii) DUSP5 (PDB structure = 2G6Z),<sup>35</sup> and (iii) FN1 (PDB structure = 6XAX).<sup>36</sup> To specify the significance of affinity binding score, we used drugs with high TF-IDF value as positive controls, namely 16-hydroxytryptolide, doxorubicin, and vorinostat. Evaluation of molecular docking simulations revealed that the following drugs 4-demethoxydaunorubicin hydrochloride targeting CD55 (with a binding affinity of  $-7.4$  kcal/mol), BRD-A60245366 targeting DUSP5 (with a binding affinity of  $-6.5$  kcal/mol), sinensetin, AS605240, and TWS119 targeting FN1 (with a binding affinity of  $-6.5$  kcal/mol,  $-6.7$  kcal/mol and  $-7.5$  kcal/mol respectively) and ER 27319 maleate targeting DUSP5 and FN1 (with binding affinities of  $-6.7$  kcal/mol and  $-7.0$  kcal/mol, respectively) have a significant docking score (kcal/mol) than the controls studied. Therefore, these six drugs represented novel drugs for the treatment of PTC in this study, and their efficiency was also demonstrated by *in silico* analyses (Figure 4B).



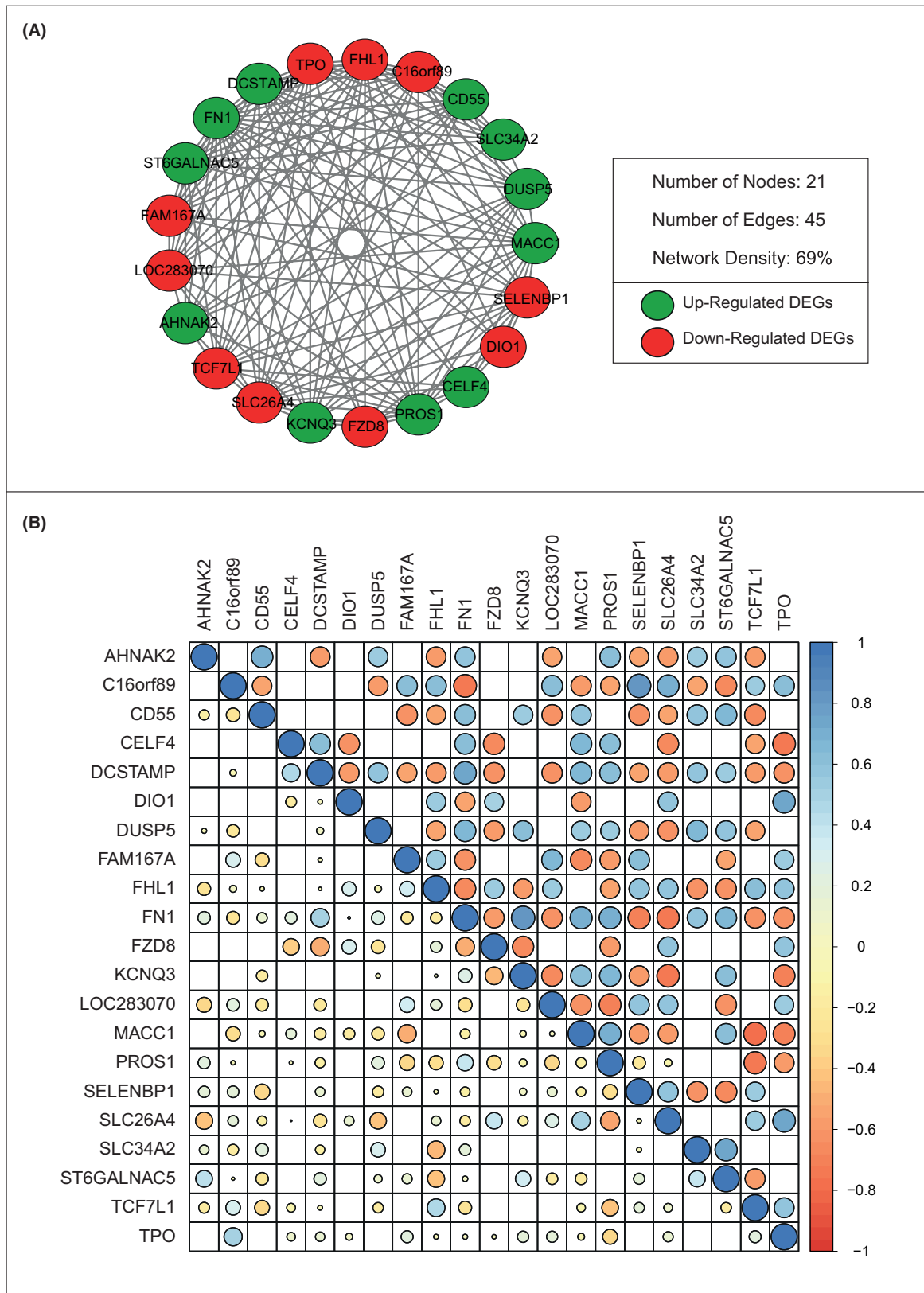
**FIGURE 1** Meta-analysis of the five transcriptome datasets associated with papillary thyroid cancer (PTC). (A) The number of differentially expressed genes (DEGs) identified for each dataset. (B) Venn diagram showing the common DEGs between all transcriptome datasets. (C) The pathway over-representation analysis of the common DEGs.

## 4 | DISCUSSION

Thyroid cancer is the most common form of endocrine cancer and PTC is the most common form of thyroid cancer. Since the true risk factors for the disease are unknown, there is still a need to identify prognostic biomarkers and improve existing treatment strategies to improve quality of life and reduce the risk of mortality. In addition, we must remember that biomolecules never function alone, but work together in a complex, interconnected network. Therefore,

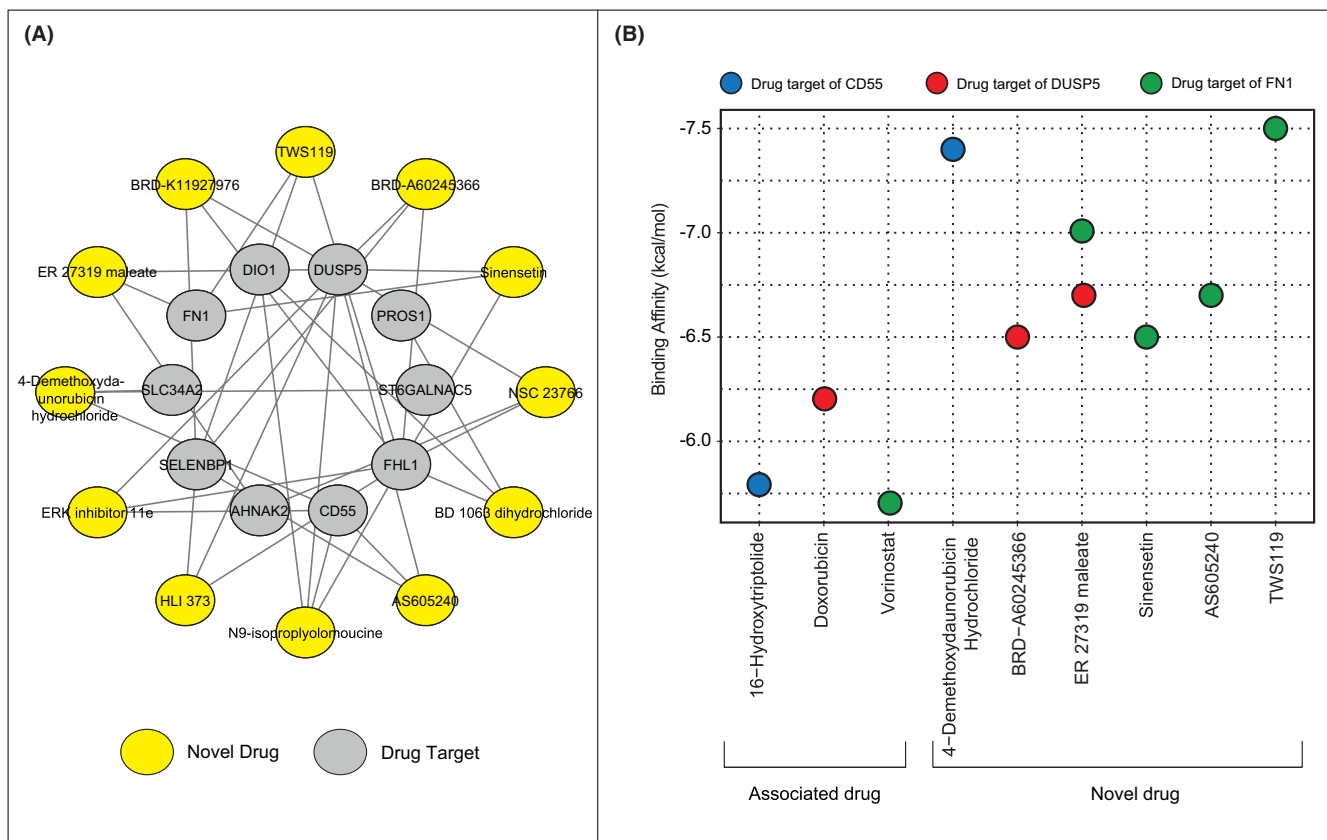
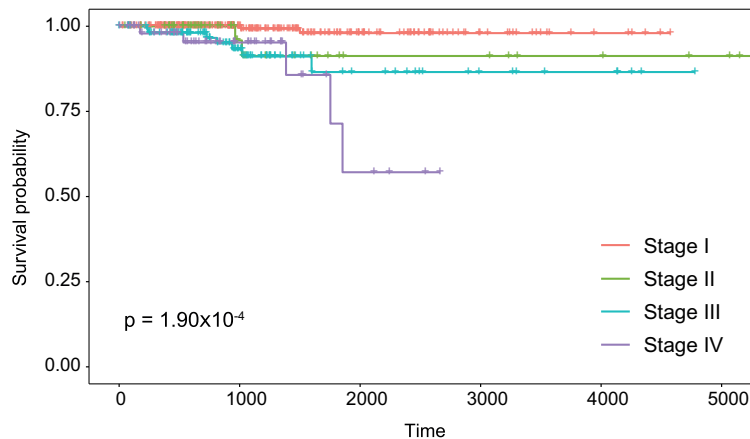
when discovering biomarkers or drugs, it is important not only to evaluate individual genes, but also to consider the system as a whole and to use the genes taking into account the interactions between them. Accordingly, in this study, we performed bioinformatics meta-analyses to identify prognostic system biomarkers and novel drug candidates for PTC.

In our analyzes based on five independent transcriptome datasets, we first attempted to detect differences between the PTC and control phenotypes (i.e. DEG analysis). To increase the reliability of



**FIGURE 2** The PTC-module (A) Co-expressed differentially expressed genes within the PTC-module. (B) The correlation plot of the PTC-module. In the correlation plot the 'lower triangle' represents the co-expression patterns of genes in healthy state, whereas the 'upper triangle' represents the co-expression patterns of genes in diseased state. Correlation coefficient is used to quantify the co-expression pattern.

**FIGURE 3** Prognostic performance analysis of the PTC-module. Kaplan–Meier diagram for estimating patient survival based on stage information with a p-value.



**FIGURE 4** The novel drug candidates and molecular docking simulations. (A) The 12 novel drug candidates and their targets. The novel drug candidates are shown in yellow, while the targets are shown in grey. (B) The bubble plot shows the molecular docking scores of the novel drug candidates. Only the novel drugs that have a significant docking score compared to the drugs that were associated with the disease (i.e. the controls) are shown.

the study and the robustness of the DEG results, we integrated the resulting DEGs in each dataset and focused only on the common 543 DEGs. Analysis of the overrepresentation of common DEGs revealed that several signalling pathways are already associated with thyroid cancer. Indeed, overexpression of TGF- $\beta$ 1 has been reported to promote PTC cell invasion and migration.<sup>37</sup> In thyroid carcinogenesis, it became clear that the PI3K/Akt pathway plays an essential role.<sup>38</sup> In the progression of PTC, deregulation of focal adhesion kinase (FAK) is observed at both post-transcriptional and translational

levels.<sup>39</sup> Collagen is one of the most abundantly expressed proteins in humans, and collagens are involved at every level of cancer development and establish interactions with cancer cells.<sup>40</sup> One of the recent studies has shown that collagen family genes (especially COL18A1 overexpression) are associated with poor prognosis and can be used as prognostic biomarkers for PTC.<sup>41</sup>

After finding co-expressed gene pairs, we created a differential co-expression network and identified differential co-expressed gene modules. According to certain criteria, one module, called the

'PTC-module,' was found to be highly clustered and also to have high prognostic significance. When we evaluate the compromising 21 nodes of this PTC-module, we find that several of these genes have already been individually associated with thyroid cancer, which further strengthens our confidence in our observations when we consider them as systems biomarkers. The following genes: PROS1,<sup>42</sup> DCSTAMP,<sup>43</sup> DUSP5,<sup>44</sup> AHNAK2,<sup>45</sup> SLC34A2,<sup>46</sup> FN1,<sup>47</sup> C16orf89,<sup>48</sup> FHL1,<sup>49</sup> TPO,<sup>50</sup> DIO1<sup>51</sup> has been associated with thyroid cancer and/or PTC.

In this study, to find novel treatment strategies for PTC, we used PTC-module nodes as proxies and performed drug repurposing analysis, revealed the novelty of drugs by text mining and demonstrated their efficiency by molecular docking simulations. We identified a total of six novel drug candidates (4-demethoxydaunorubicin hydrochloride, AS605240, BRD-A60245366, ER 27319 maleate, sinensetin and TWS119) for the treatment of PTC, whose potential efficacy was also demonstrated by molecular docking simulations. 4-Demethoxydaunorubicin hydrochloride, also known as idarubicin hydrochloride, is a type of anthracycline antibiotic used to treat acute myeloid leukaemia. To execute its anticancer effects, the drug inhibits topoisomerase II, which this resulting in impaired DNA replication.<sup>52</sup> A PI3K $\gamma$ -selective inhibitor, AS605240, has shown associations with glomerulonephritis, rheumatoid arthritis,<sup>53</sup> and Alzheimer's disease<sup>54</sup> in mouse models. BRD-A60245366, also known as AS601245, is a JNK inhibitor and has anti-inflammatory properties. It has been associated with colon cancer. It has been reported that the combined treatment of BRD-A60245366 and clofibrate has an effect on colon cancer cell lines (i.e. induces cell responses).<sup>55</sup> ER 27319 maleate is a selective inhibitor of Syk kinase and has been associated with head and neck cancer.<sup>56</sup> Sinensetin has multiple pharmacological activities, including antineoplastic, anti-inflammatory and antimicrobial activities. It has been associated with gastric, colon, and breast cancer.<sup>57</sup> TWS119 is a potent inhibitor of GSK-3 $\beta$  and it has been reported that TWS119 increases the replication and cytolytic activity of human  $\gamma\delta$ T cells against human colon cancer cells. These  $\gamma\delta$ T cells represent unique T cell subpopulations and can be used in cancer immunotherapy. Therefore, TWS119 has been used as a complementary agent to enhance  $\gamma\delta$ T-cell-based immunotherapy for colorectal cancer.<sup>58</sup>

In conclusion, the distribution of possible significant changes in gene pairs during the transition from a normal to a tumour phenotype is of great importance. Moreover, the evaluation of the modular structures around these co-expressed gene pairs will provide us with much more effective results in the development of diagnostic, prognostic biomarkers, and/or treatment strategies. Therefore, in this study, we investigated the different co-expression profiles between PTC and control phenotypes. Based on this information, we proposed a module (i.e. the PTC-module) with a high prognostic capacity and suggested it as a system biomarker. In addition, six novel drug candidates for PTC were presented in this study, which deserve to be investigated with experimental studies. The benefits of integrative bioinformatics approaches are also highlighted to develop a network medicine approach for thyroid cancer prognosis and treatment in the future.

## AUTHOR CONTRIBUTIONS

**Medi Kori:** Data curation (equal); formal analysis (equal); methodology (equal); writing – original draft (equal); writing – review and editing (equal). **Kubra Temiz:** Data curation (equal); formal analysis (equal); writing – original draft (equal). **Esra Gov:** Conceptualization (equal); methodology (equal); supervision (equal); writing – review and editing (equal).

## ACKNOWLEDGEMENTS

This study is supported by the Health Institutes of Turkey (TÜSEB) under project 2019-TA-01/3436.

## CONFLICT OF INTEREST STATEMENT

The authors confirm that there are no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Esra Gov  <https://orcid.org/0000-0002-5256-4778>

## REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72(1):7-33. doi:10.3322/caac.21708
2. Coca-Pelaz A, Shah JP, Hernandez-Prera JC, et al. Papillary thyroid cancer-aggressive variants and impact on management: a narrative review. *Adv Ther.* 2020;37(7):3112-3128. doi:10.1007/s12325-020-01391-1
3. Kori M, Gov E. Bioinformatics prediction and machine learning on gene expression data identifies novel gene candidates in gastric cancer. *Genes.* 2022;13(12):2233.
4. Kori M, Arga K. HPV16 status predicts potential protein biomarkers and therapeutics in head and neck squamous cell carcinoma. *Virology.* 2023;582:90-99. doi:10.1016/j.virol.2023.03.013
5. Gulfidan G, Soylu M, Demirel D, et al. Systems biomarkers for papillary thyroid cancer prognosis and treatment through multi-Omics networks. *Arch Biochem Biophys.* 2022;715:109085. doi:10.1016/j.abb.2021.109085
6. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12(1):56-68. doi:10.1038/nrg2918
7. Savino A, Provero P, Poli V. Differential Co-expression analyses allow the identification of critical Signalling pathways altered during tumour transformation and progression. *Int J Mol Sci.* 2020;21(24):9461. doi:10.3390/ijms21249461
8. Tang X, Huang X, Wang D, et al. Identifying gene modules of thyroid cancer associated with pathological stage by weighted gene Co-expression network analysis. *Gene.* 2019;704:142-148. doi:10.1016/j.gene.2019.04.017
9. Wang B, Jiang W, Zheng X, Han Y, Liu R. Research on a weighted gene Co-expression network analysis method for mining pathogenic genes in thyroid cancer. *PLoS One.* 2022;17(8):e0272403. doi:10.1371/journal.pone.0272403
10. Hou J, Ye X, Li C, Wang Y. K-module algorithm: an additional step to improve the clustering results of WGCNA Co-expression networks. *Genes.* 2021;12(1):87. doi:10.3390/genes12010087
11. Gov E, Arga KY. Differential Co-expression analysis reveals a novel prognostic gene module in ovarian cancer. *Sci Rep.* 2017;7(1):4996.

12. Kori M, Gov E, Arga KY. novel genomic biomarker candidates for cervical cancer As identified by differential Co-expression network analysis. *Integr Biol*. 2019;23(5):261-273. doi:10.1089/omi.2019.0025
13. Gov E. Co-expressed functional module-related genes in ovarian cancer stem cells represent novel prognostic biomarkers in ovarian cancer. *Syst Biol Reprod Med*. 2020;66(4):255-266. doi:10.1080/19396368.2020.1759730
14. Demirtas TY, Rahman MR, Yurtsever MC, Gov E. forecasting gastric cancer diagnosis, prognosis, and drug repurposing with novel gene expression signatures. *Integr Biol*. 2022;26(1):64-74.
15. Tarabichi M, Saiselet M, Trésallet C, et al. Revisiting the transcriptional analysis of primary tumours and associated nodal metastases with enhanced biological and statistical controls: application to thyroid cancer. *Br J Cancer*. 2015;112(10):1665-1674. doi:10.1038/bjc.2014.665
16. Handkiewicz-Junak D, Swierniak M, Rusinek D, et al. Gene signature of the post-Chernobyl papillary thyroid cancer. *Eur J Nucl Med Mol Imaging*. 2016;43(7):1267-1277. doi:10.1007/s00259-015-3303-3
17. Tomás G, Tarabichi M, Gacquer D, et al. A general method to derive robust organ-specific gene expression-based differentiation indices: application to thyroid cancer diagnostic. *Oncogene*. 2012;31(41):4490-4498. doi:10.1038/nc.2011.626
18. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: Archive for Functional Genomics Data Sets—Update. *Nucleic Acids Res*. 2013;41(Database issue):D991-995:D991-D995. doi:10.1093/nar/gks1193
19. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015;19(1A):A68-A77. doi:10.5114/wo.2014.47136
20. Kori M, Cig D, Arga KY, Kasavi C. Multiomics data integration identifies new molecular signatures for abdominal aortic aneurysm and aortic occlusive disease: implications for early diagnosis, prognosis, and therapeutic targets. *Omics*. 2022;26(5):290-304. doi:10.1089/omi.2022.0021
21. Unal U, Comertpay B, Demirtas TY, Gov E. Drug repurposing for rheumatoid arthritis: identification of new drug candidates via bioinformatics and text mining analysis. *Autoimmunity*. 2022;55(3):147-156.
22. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide Array data based on variance and bias. *Bioinformatics (Oxford, England)*. 2003;19(2):185-193. doi:10.1093/bioinformatics/19.2.185
23. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy—Analysis of Affymetrix GeneChip Data at the Probe Level. *Bioinformatics (Oxford, England)*. 2004;20(3):307-315. doi:10.1093/bioinformatics/btg405
24. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-121. doi:10.1038/nmeth.3252
25. Ritchie ME, Phipson B, Di W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
26. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB—a Database for Integrating Human Functional Interaction Networks. *Nucleic Acids Res*. 2009;37(Database issue):D623-628:D623-D628. doi:10.1093/nar/gkn698
27. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30. doi:10.1093/nar/28.1.27
28. Fabregat A, Sidiropoulos K, Garapati P, et al. The Reactome pathway knowledgebase. *Nucleic Acids Res*. 2016;44(D1):D481-D487. doi:10.1093/nar/gkv1351
29. Duan Q, Reid SP, Clark NR, et al. L1000CDS2: LINCS L1000 characteristic direction signatures search engine. *Npj Systems Biology and Applications*. 2016;2(1):1-12. doi:10.1038/npijsba.2016.15
30. Cock PJA, Antao T, Chang JT, et al. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics*. 2009;25(11):1422-1423. doi:10.1093/bioinformatics/btp163
31. Berman HM, Narayanan BC, Di Costanzo L, et al. Trendspotting in the protein data Bank. *FEBS Lett*. 2013;587(8):1036-1045. doi:10.1016/j.febslet.2012.12.029
32. Kim S, Chen J, Cheng T, et al. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res*. 2019;47(D1):D1102-D1109. doi:10.1093/nar/gky1033
33. Trott O, Olson AJ. AutoDock Vina: improving the Speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31(2):455-461. doi:10.1002/jcc.21334
34. Williams P, Chaudhry Y, Goodfellow IG, et al. Mapping CD55 function. *J Biol Chem*. 2003;278(12):10691-10696. doi:10.1074/jbc.M212561200
35. Jeong DG, Cho YH, Yoon T-S, Kim JH, Ryu SE, Kim SJ. Crystal structure of the catalytic domain of human DUSP5, a dual specificity MAP kinase protein phosphatase. *Proteins: Structure, Function, and Bioinformatics*. 2006;66(1):253-258. doi:10.1002/prot.21224
36. Bencharit S, Cui CB, Siddiqui A, et al. Structural insights into fibronectin type III domain-mediated signaling. *J Mol Biol*. 2007;367(2):303-309. doi:10.1016/j.jmb.2006.10.017
37. Xie J, Liu Y, Xiaolong D, Ying W. TGF-B1 promotes the invasion and migration of papillary thyroid carcinoma cells by inhibiting the expression of LncRNA-NEF. *Oncol Lett*. 2019;17(3):3125-3132. doi:10.3892/ol.2019.9947
38. Robbins HL, Hague A. The PI3K/Akt pathway in tumors of endocrine tissues. *Front Endocrinol*. 2015;6:188. doi:10.3389/fendo.2015.00188
39. Ignjatović VB, Janković JR, Miljuš JV, et al. Focal adhesion kinase splicing and protein activation in papillary thyroid carcinoma progression. *Histochem Cell Biol*. 2022;157(2):183-194. doi:10.1007/s00418-021-02056-y
40. Shi R, Zhang Z, Zhu A, et al. Targeting type I collagen for cancer treatment. *Int J Cancer*. 2022;151(5):665-683. doi:10.1002/ijc.33985
41. Niu J, Guo W, Chen Y-Z, Jiang N. Identification of the collagen family as prognostic biomarkers in papillary thyroid carcinoma. *Endocrine*. 2022;78(3):491-506. doi:10.1007/s12020-022-03175-9
42. Wang J, Lei M, Zhijie X. Aberrant expression of PROS1 correlates with human papillary thyroid cancer progression. *PeerJ*. 2021;9:e11813. doi:10.7717/peerj.11813
43. Lee J-J, Hsu Y-C, Huang W-C, Cheng S-P. Upregulation of Dendrocyte-expressed seven transmembrane protein is associated with unfavorable outcomes in differentiated thyroid cancer. *Endocrine*. 2023;81:513-520. doi:10.1007/s12020-023-03364-0
44. Buffet C, Hecale-Perlemoine K, Bricaire L, et al. DUSP5 and DUSP6, two ERK specific phosphatases, are markers of a higher MAPK signaling activation in BRAF mutated thyroid cancers. *PLoS One*. 2017;12(9):e0184861. doi:10.1371/journal.pone.0184861
45. Ye R, Liu D, Guan H, et al. AHNK2 promotes thyroid carcinoma progression by activating the NF-KB pathway. *Life Sci*. 2021;286:120032.
46. He J, Zhou M, Li X, et al. SLC34A2 simultaneously promotes papillary thyroid carcinoma growth and invasion through distinct mechanisms. *Oncogene*. 2020;39(13):2658-2675. doi:10.1038/s41388-020-1181-z
47. Chen C, Shen Z. FN1 promotes thyroid carcinoma cell proliferation and metastasis by activating the NF-Kb pathway. *Protein Pept Lett*. 2023;30(1):54-64. doi:10.2174/0929866530666221019162943
48. Afink GB, Veenboer G, de Randamie J, et al. Initial characterization of C16orf89, a novel thyroid-specific gene. *Thyroid: Official Journal of the American Thyroid Association*. 2010;20(7):811-821. doi:10.1089/thy.2009.0366

49. Wei X, Zhang H. Four and a half LIM domains protein 1 can be as a double-edged sword in cancer progression. *Cancer Biol Med*. 2020;17(2):270-281. doi:[10.20892/j.issn.2095-3941.2019.0420](https://doi.org/10.20892/j.issn.2095-3941.2019.0420)
50. Caballero Y, López-Tomassetti EM, Favre J, Santana JR, Cabrera JJ, Hernández JR. The Value of Thyroperoxidase as a Prognostic Factor for Differentiated Thyroid Cancer—a Long-Term Follow-up Study. *Thyroid Res*. 2015;8:12. doi:[10.1186/s13044-015-0022-6](https://doi.org/10.1186/s13044-015-0022-6)
51. AlRasheed MM, AlAnzi A, AlShalhoub R, Abanmy N, Bakheet D. A study of the role of DIO1 and DIO2 polymorphism in thyroid cancer and drug response to therapy in the Saudi population. *Saudi Pharmaceutical Journal: SPJ: The Official Publication of the Saudi Pharmaceutical Society*. 2019;27(6):841-845. doi:[10.1016/j.jsps.2019.05.005](https://doi.org/10.1016/j.jsps.2019.05.005)
52. Sekine L, Morais VD, Lima KM, Onsten TGH, Ziegelmann PK, Ribeiro RA. Conventional and high-dose daunorubicin and idarubicin in acute myeloid leukaemia remission induction treatment: a mixed treatment comparison meta-analysis of 7258 patients. *Hematol Oncol*. 2015;33:212-219. doi:[10.1002/hon.2173](https://doi.org/10.1002/hon.2173)
53. Kong D, Yamori T. Phosphatidylinositol 3-kinase inhibitors: promising drug candidates for cancer therapy. *Cancer Sci*. 2008;99(9):1734-1740. doi:[10.1111/j.1349-7006.2008.00891.x](https://doi.org/10.1111/j.1349-7006.2008.00891.x)
54. Alluri R, Ambati SR, Routhu K, Kopalli SR, Koppula S. Phosphoinositide 3-kinase inhibitor AS605240 ameliorates Streptozotocin-induced Alzheimer's disease like sporadic dementia in experimental rats. *EXCLI J*. 2020;19:71-85. doi:[10.17179/excli2019-1997](https://doi.org/10.17179/excli2019-1997)
55. Cerbone A, Toaldo C, Pizzimenti S, et al. AS601245, an anti-inflammatory JNK inhibitor, and Clofibrate have a synergistic effect in inducing cell responses and in affecting the gene expression profile in CaCo-2 colon cancer cells. *PPAR Research*. 2012;2012:269751. doi:[10.1155/2012/269751](https://doi.org/10.1155/2012/269751)
56. Black M, Ghasemi F, Sun RX, et al. Spleen tyrosine kinase expression is correlated with human papillomavirus in head and neck cancer. *Oral Oncol*. 2020;101:104529. doi:[10.1016/j.oraloncology.2019.104529](https://doi.org/10.1016/j.oraloncology.2019.104529)
57. Patel K, Patel DK. Therapeutic effectiveness of Sinensetin against cancer and other human complications: a review of biological potential and pharmacological activities. *Cardiovasc Hematol Disord Drug Targets*. 2022;22:144-154. doi:[10.2174/1871529X23666221207121955](https://doi.org/10.2174/1871529X23666221207121955)
58. Chen Y-Q, Zheng L, Aldarouish M, et al. Wnt pathway activator TWS119 enhances the proliferation and Cytolytic activity of human  $\Gamma\delta$ T cells against colon cancer. *Exp Cell Res*. 2018;362(1):63-71. doi:[10.1016/j.yexcr.2017.11.003](https://doi.org/10.1016/j.yexcr.2017.11.003)

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Kori M, Temiz K, Gov E. Network medicine approaches for identification of novel prognostic systems biomarkers and drug candidates for papillary thyroid carcinoma. *J Cell Mol Med*. 2023;27:4171-4180. doi:[10.1111/jcmm.18002](https://doi.org/10.1111/jcmm.18002)