



Integrative Analysis of MicroRNA and Gene Interactions for Revealing Candidate Signatures in Prostate Cancer

Jingchao Wei¹, Yinghao Yin¹, Qiancheng Deng², Jun Zhou¹, Yong Wang³, Guangming Yin^{1*}, Jianfu Yang¹ and Yuxin Tang^{1*}

¹ Department of Urology, The Third Xiangya Hospital of Central South University, Changsha, China, ² Department of Dermatology, The Second Xiangya Hospital of Central South University, Changsha, China, ³ Department of Urology, Hunan Provincial People's Hospital, Changsha, China

OPEN ACCESS

Edited by:

Chandan Kumar-Sinha,
University of Michigan, United States

Reviewed by:

Andi Cani,
University of Michigan, United States

Marco Maugeri,
University of Skövde, Sweden

*Correspondence:

Guangming Yin
yingm75@126.com
Yuxin Tang
mmcct@126.com

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 27 November 2019

Accepted: 13 February 2020

Published: 27 February 2020

Citation:

Wei J, Yin Y, Deng Q, Zhou J, Wang Y, Yin G, Yang J and Tang Y (2020) Integrative Analysis of MicroRNA and Gene Interactions for Revealing Candidate Signatures in Prostate Cancer. *Front. Genet.* 11:176. doi: 10.3389/fgene.2020.00176

MicroRNA (miRNA)–gene interactions are well-recognized as involved in the progression of almost all cancer types including prostate cancer, which is one of the most common cancers in men. This study explored the significantly dysregulated genes and miRNAs and elucidated the potential miRNA–gene regulatory network in prostate cancer. Integrative analysis of prostate cancer and normal prostate transcriptomic data in The Cancer Genome Atlas dataset was conducted using both differential expression analysis and weighted correlation network analysis (WGCNA). Thirteen genes (*RRM2*, *ORC6*, *CDC45*, *CDKN2A*, *E2F2*, *MYBL2*, *CCNB2*, *PLK1*, *FOXM1*, *CDC25C*, *PKMYT1*, *GTSE1*, and *CDC20*) were potentially correlated with prostate cancer based on functional enrichment analyses. MiRNAs targeting these genes were predicted and eight miRNAs were intersections between those miRNAs and the hub miRNAs obtained from miRNA WGCNA analysis. Three genes (*E2F2*, *RRM2*, and *PKMYT1*) and four miRNAs (*hsa-mir-17-5p*, *hsa-mir-20a-5p*, *hsa-mir-92a-3p*, and *hsa-mir-93-5p*) were key factors according to the interaction network. *RRM2* and *PKMYT1* were significantly related to survival. These findings partially elucidated the dysregulation of gene expressions in prostate cancer. Efficient manipulations of the miRNA–gene interactions in prostate cancer may be exploited as promising therapeutics.

Keywords: prostate cancer, signature, microRNA, WGCNA, TCGA

INTRODUCTION

Prostate cancer ranks as one of the most common malignancies in men (Attard et al., 2016). High-risk prostate cancer is aggressive and can lead to poor outcomes at advanced stages and metastasis (Litwin and Tan, 2017). A tremendous number of investigations have been conducted with the goal of determining the mechanisms underlying the oncogenesis and development of prostate cancer. Significant findings including androgen receptor mechanism have been well-established (Culig and Santer, 2014). However, the detailed underlying mechanism is still far from being clear. To improve the management of prostate cancer, the identification of novel signatures is of vital importance for the prediction of prognosis and targeted therapy of prostate cancer patients.

MicroRNAs (miRNAs) are small, non-protein-coding RNA (usually 22 nucleotides) molecules, which function in RNA silencing and post-transcriptional regulation of target gene expression (Rupaimoole and Slack, 2017). Accumulating evidence has shown that miRNAs are involved in nearly all cancer types, acting as tumor oncogenes or suppressors and taking part in cell proliferation, differentiation, and metastasis (Winbanks et al., 2014; Vorvis et al., 2016). It is well-established that a large number of specific miRNAs are involved in the carcinogenesis and development of tumors by regulating the expressions of their targeted mRNAs (Inui et al., 2010). With advancements in biological and clinical research, miRNA-mRNA interactions have been widely demonstrated to regulate the complicated molecular mechanisms underlying oncogenesis, development, invasion, and metastasis of tumors (Duell et al., 2017; Han et al., 2017). The Cancer Genome Atlas (TCGA) database has stored numerous genomic and gene expression profiles for various types of tumors. Many genes and signaling pathways correlated with cancer have been predicted based on analyses of these data, which have provided valuable guidance for subsequent molecular biology validation (The TCGA Legacy, 2018).

This study investigated the differentially expressed genes (DEGs) and hub miRNAs in prostate cancer and screened signatures of prostate cancer. Integrative analysis was performed using gene and miRNA expression data of prostate cancer downloaded from TCGA database to identify DEGs and miRNAs in prostate cancer when compared with adjacent normal tissue samples. Functional enrichment analyses were subsequently conducted. Then the mRNA-miRNA interaction network was constructed. Three genes and four miRNAs were found to be hub factors of prostate cancer and the prognostic significance was investigated.

MATERIALS AND METHODS

Analysis of Differentially Expressed Genes Between Prostate Cancer and Normal Samples in TCGA Prostate Adenocarcinoma (PRAD) Datasets

Transcriptome raw count data of TCGA PRAD were obtained from GDC Data Portal using R TCGAbiolinks package (Colaprico et al., 2016). Clinical data of TCGA PRAD subjects were obtained from UCSC XENA browser.¹ Samples of primary solid tumor and solid normal tissue were picked out for analysis, and paraffin embedded samples as well as metastatic samples were removed. Ensemble identifiers were annotated with gene names using R Annotables package (genome version GRCh38). Identifiers with no annotations were removed. For duplicated identifiers corresponding to one gene name, the average of all count values was calculated and rounded to be taken as the count of that gene. Analysis of DEGs between prostate cancer and normal tissues were performed using R DESeq2 package (Love et al., 2014). The phenotype of tumor or normal

sample was used as grouping variable and standard differential expression analysis was conducted. The normalized expression matrix was obtained through variance stabilizing transformation of raw count data. As the DEGs would be taken as input in the subsequent analysis, a less stringent criterion was used to identify DEGs: adjusted p -value < 0.05 and absolute value of fold change ≥ 2 . Principal component analysis (PCA) was performed based on the normalized expression data, using `prcomp` function in R. `Pheatmap` and `ggplot2` packages were used to plot gene expression boxplot, volcano plot and heatmap.

Weighted Correlation Network Analysis of DEGs and the Identification of Modules Associated With Prostate Cancer

The DEGs obtained by DESeq2 were used to perform weighted correlation network analysis (WGCNA) to get the interesting gene modules based on the gene expressions and patient traits using R WGCNA package (Langfelder and Horvath, 2008). Before the one-step network construction and module detection, a soft thresholding power value was calculated to produce a scale-free network topology. Then one-step network construction and detection of consensus modules were executed. The network type was set as signed. Correlations between clinical traits (tumoral or normal) and each module were calculated and preliminary Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted for genes in several top trait-correlated modules to determine the module that made the most sense. Black module was speculated to be a candidate tumor-driving module since preliminary KEGG analyses revealed some interesting pathways potentially related with tumor biology.

Functional Enrichment Analyses of Genes in Significantly Correlated Modules

To examine the biological functions of the genes in black modules, we performed functional enrichment analyses using KEGG, Gene Ontology (GO), and REACTOME Pathway databases. KEGG pathway enrichment analysis was conducted using the R package `clusterProfiler` (Yu et al., 2012). Genes in the black module were annotated with R package `Annotables` and all DEGs obtained by DESeq2 were used as background. Twelve pathways were significantly enriched (adjusted p -value < 0.05), and genes in five interesting pathways among the 12 pathways were picked out for subsequent GO biological process analysis, which was performed using `g:GOST` tool in `gProfiler`.² `gProfiler` is a web server for functional enrichment analysis and conversions of gene lists. The tool `g:GOST` in `gProfiler` performs functional enrichment analysis on input gene list (Raudvere et al., 2019). Among all significantly enriched biological processes, the top 20 were selected for investigation, and the associated genes in all 20 interesting terms were used for subsequent REACTOME

¹<https://xena.ucsc.edu>

²<https://biit.cs.ut.ee/gprofiler>

Pathway enrichment analysis using the REACTOME Pathway analysis tool.³ Among all significantly enriched REACTOME pathways, we studied the top 20 and identified 13 genes in nine pathways that were potentially prostate cancer gene drivers. Enrichplot, ggplot2, and pheatmap packages were used for plotting.

Weighted Correlation Network Analysis of miRNAs and Identification of Hub miRNAs Associated With Prostate Cancer

Normalized transcriptome data of miRNA [$\log_2(\text{total_RPM} + 1)$ transformed] in prostate cancer were obtained from UCSC XENA¹. Likewise, samples of primary solid tumor and solid normal tissue were selected for the analysis, and paraffin embedded samples and metastatic samples were removed. MiRNA identifiers were annotated with detailed names using the R MiRBaseConverter package (Xu et al., 2018) (target version V22). Identifiers with no annotations were removed. For duplicate identifiers corresponding to one miRNA name, the average of all values was calculated to represent the expression value for that miRNA. Then the processed miRNA expression data were used to perform WGCNA to obtain interesting miRNA modules correlated with patient traits (tumoral or normal). The pattern for constructing one-step network and detecting consensus miRNA modules was similar to that of the aforementioned genes. Correlations between clinical traits and each module were calculated and the turquoise module was the most correlative module with prostate cancer. Then top-ranked miRNAs that had maximum connectivity with other miRNAs were determined by the signedKME function in the turquoise module.

Construction of the Gene–miRNA Interaction Network and Identification of Prostate Cancer Signatures

The MiRNet tool,⁴ which is an integrated platform linking miRNAs, was used to collect all miRNAs that were predicted or validated to target the 13 candidate gene drivers. The miRNA–gene interaction data of miRNet were collected from miRTarBase v7.0, TarBase v7.0 and miRecords, catalogs that predict and validate miRNA–gene interactions (Xiao et al., 2009; Vlachos et al., 2015; Chou et al., 2018). The validated miRNA–gene interactions are manually curated and experimentally validated, and the predicted miRNA–gene interactions are produced by established miRNA target prediction programs. MiRNAs targeting at least two genes were screened to identify intersections with the hub miRNAs obtained by WGCNA analysis. Gene–miRNA interactions were plotted using Cytoscape software (Shannon et al., 2003).

Kaplan–Meier Survival Analyses and Receiver Operating Characteristic Curves of the Three Gene Signatures

To investigate the clinical significance of the three gene signatures, we performed Kaplan–Meier survival analyses and receiver operating characteristic (ROC) curves for survival. Survival data of TCGA Prostate Adenocarcinoma were obtained from the UCSC XENA platform, and the gene expression data used in this section were the aforementioned normalized expression matrix obtained by DESeq2 package. Kaplan–Meier survival analyses were conducted utilizing survfit function in the R Survival package. Survival ROC curve analyses for survival prediction were performed using roc function in the R proc package based on the miRNA expression data obtained from UCSC XENA (Robin et al., 2011). Area under curve (AUC) was calculated to measure how well the model could distinguish between survival and non-survival classes.

Validation of Expression of Four miRNA Signatures in Prostate Cancer Using Gene Expression Omnibus Datasets

To analyze the expression of four miRNA signatures in current high-throughput projects on prostate cancer, we used dbDEMC (version 2.0⁵), which is an integrated database designed to store and display differentially expressed miRNAs in human cancers detected by high-throughput methods (Yang et al., 2017). R limma package was used to screen the differentially expressed miRNAs from cancer compared with the normal state (Ritchie et al., 2015). Fold changes and *p*-values were plotted using ggplot2 package.

Data Processing and Statistical Analysis

Data sources were mentioned above and R Programming software was obtained from the R official website (version 3.6.0⁶). Rstudio software was obtained from the official website.⁷ $p < 0.05$ was considered statistically significant.

RESULTS

Analysis of Differentially Expressed Genes (DEGs) Between Prostate Cancer and Normal Samples in TCGA Datasets

A total of 481 primary solid tumor samples and 51 solid normal samples with clinical data were screened after paraffin embedded samples and metastatic samples were removed. The DESeq2 package estimates variance-mean dependence in count data of sequencing assays and tests for differential expression based on a model using the negative binomial distribution. Since raw count data are not normally distributed, it was necessary to transform the data to a normalized scale. A normalized

³<https://reactome.org>

⁴<https://www.mirnet.ca>

⁵<http://www.picb.ac.cn/dbDEMC>

⁶<https://www.r-project.org>

⁷<https://rstudio.com>

gene expression matrix was obtained using variance stabilizing transformation function of the DESeq2 package. Box plot showed expression distributions for normalized data for each sample (**Figure 1A**). Overall, the expression distributions of

raw normalized-intensities were not identical but were similar, as no sample was that different from others. The PCA plot provided insights into the similarities among samples and indicated the quality of the expression data, which can effectively

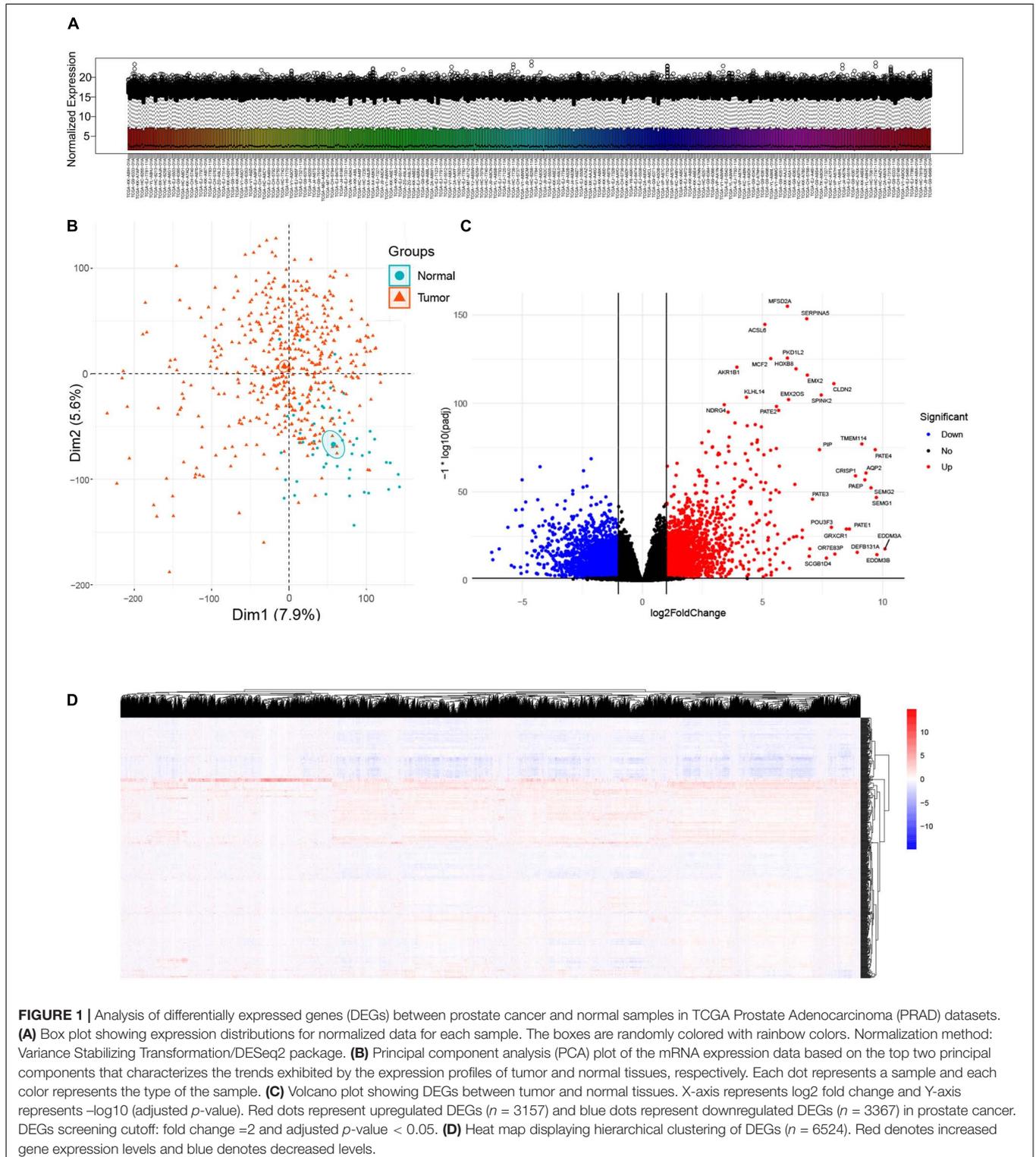


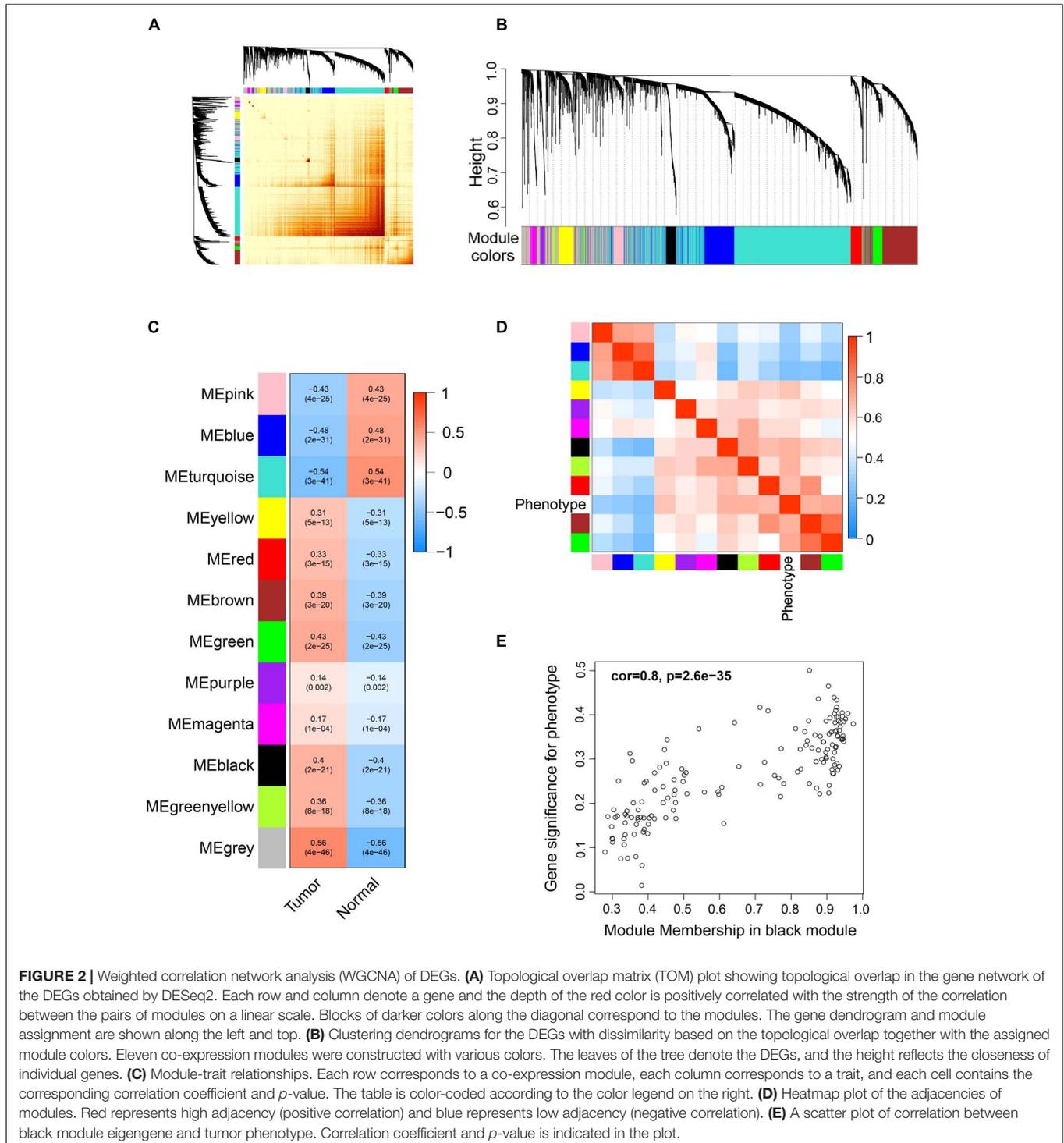
FIGURE 1 | Analysis of differentially expressed genes (DEGs) between prostate cancer and normal samples in TCGA Prostate Adenocarcinoma (PRAD) datasets. **(A)** Box plot showing expression distributions for normalized data for each sample. The boxes are randomly colored with rainbow colors. Normalization method: Variance Stabilizing Transformation/DESeq2 package. **(B)** Principal component analysis (PCA) plot of the mRNA expression data based on the top two principal components that characterizes the trends exhibited by the expression profiles of tumor and normal tissues, respectively. Each dot represents a sample and each color represents the type of the sample. **(C)** Volcano plot showing DEGs between tumor and normal tissues. X-axis represents log₂ fold change and Y-axis represents -log₁₀ (adjusted *p*-value). Red dots represent upregulated DEGs ($n = 3157$) and blue dots represent downregulated DEGs ($n = 3367$) in prostate cancer. DEGs screening cutoff: fold change = 2 and adjusted *p*-value < 0.05. **(D)** Heat map displaying hierarchical clustering of DEGs ($n = 6524$). Red denotes increased gene expression levels and blue denotes decreased levels.

discriminate low-quality samples. In this study, PCA plot based on the first two components, as they described the largest variability, showed tumor and normal samples were roughly distinguished (Figure 1B). A total of 6524 DEGs (both upregulated and downregulated genes in cancer samples compared to normal sample) were identified according to the criterion of adjusted p -value < 0.05 and absolute fold change ≥ 2

(Figure 1C). Heatmap showed the DEG expressions in all samples (Figure 1D).

WGCNA of DEGs

Differentially expressed genes were used for WGCNA analysis. WGCNA is an unsupervised analysis method that clusters genes based on their expression profiles. It is widely used to study



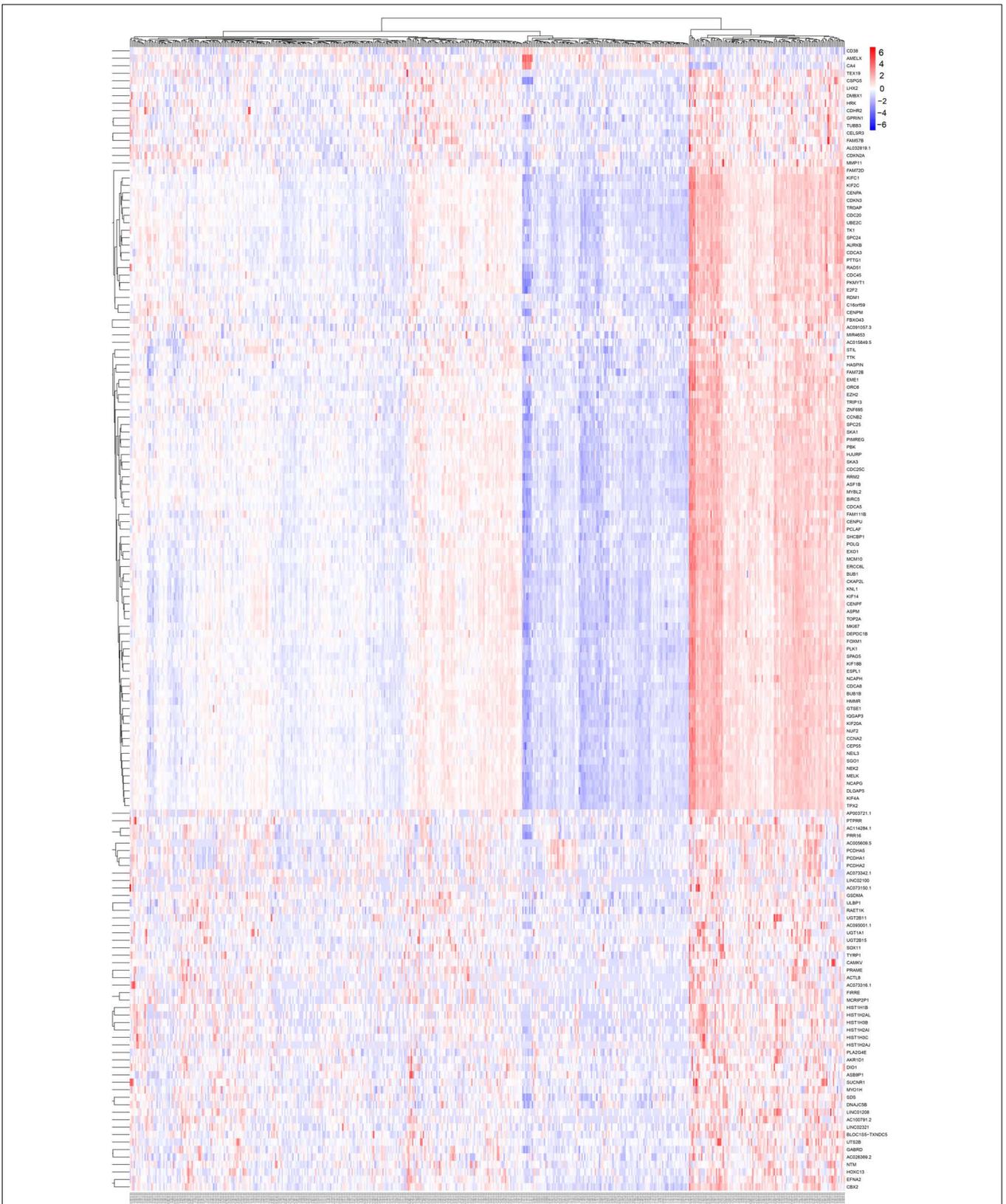
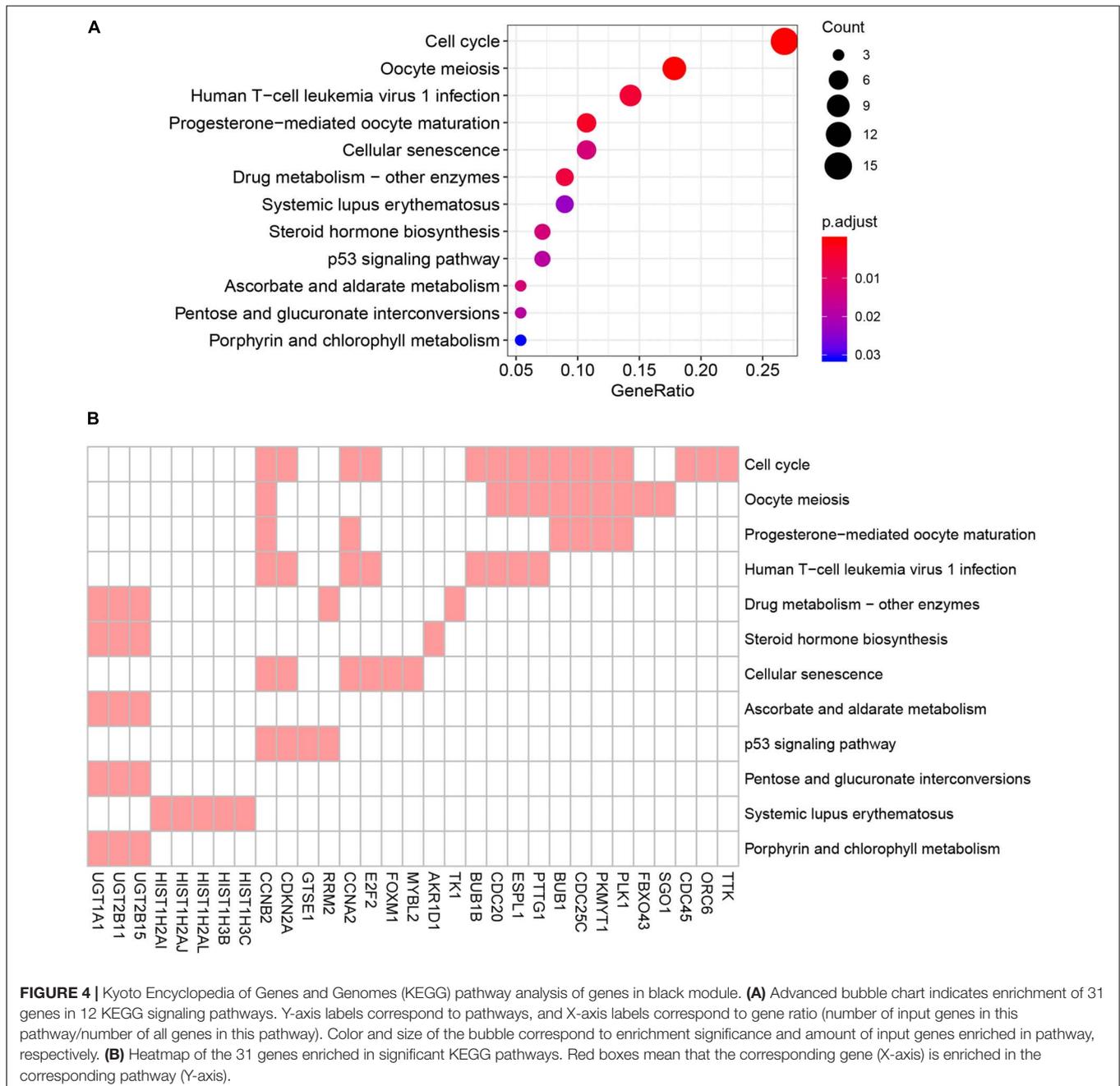


FIGURE 3 | Heatmap and cluster analysis (Euclidean distance) of the genes in black module obtained by WGCNA analysis of DEGs. Rows correspond to genes and columns correspond to samples. Red and blue colors indicate different extents of high or low expressions.

the relationships between co-expression modules (clusters) and the relationships between modules and external sample traits (Langfelder and Horvath, 2008). In our study, value 4 was chosen as the soft thresholding power value because it produced a higher similarity with a scale-free network and contributed to gene clustering. A one-step network was constructed using blockwiseModules function with the parameters: of signed Topological Overlap Matrix type and Pearson's correlation coefficient (Figure 2A). Clustering dendrograms for DEGs with dissimilarity and assigned module colors were plotted (Figure 2B), and 11 modules with a minimum module size

of 30 genes were obtained. The correlations between each module and trait (prostate cancer or normal prostate) were calculated (Figure 2C), and connection strengths (adjacencies) of the modules and traits are shown (Figure 2D). Red indicates high adjacency (positive correlation) and blue indicates low adjacency (negative correlation). Here, we identified some genes that were strongly positively correlated with prostate cancer. Among all of the modules, the green and black modules had the highest correlations with prostate cancer (excluding gray module), and preliminary KEGG pathway enrichment analyses were performed to evaluate their biological significance. Several



interesting pathways correlated with cancer were enriched by black modules, whereas analysis of the green module did not yield significant results. The black module eigengenes were highly correlated with prostate cancer (Figure 2E). Thus black module was speculated to be a potential prostate cancer-related module and was utilized for subsequent analyses.

Heatmap showed the expressions of all genes in the black module (Figure 3).

Functional Enrichment Analysis

To reduce data dimensionality, three types of enrichment analyses were utilized to obtain the most prominent genes in the

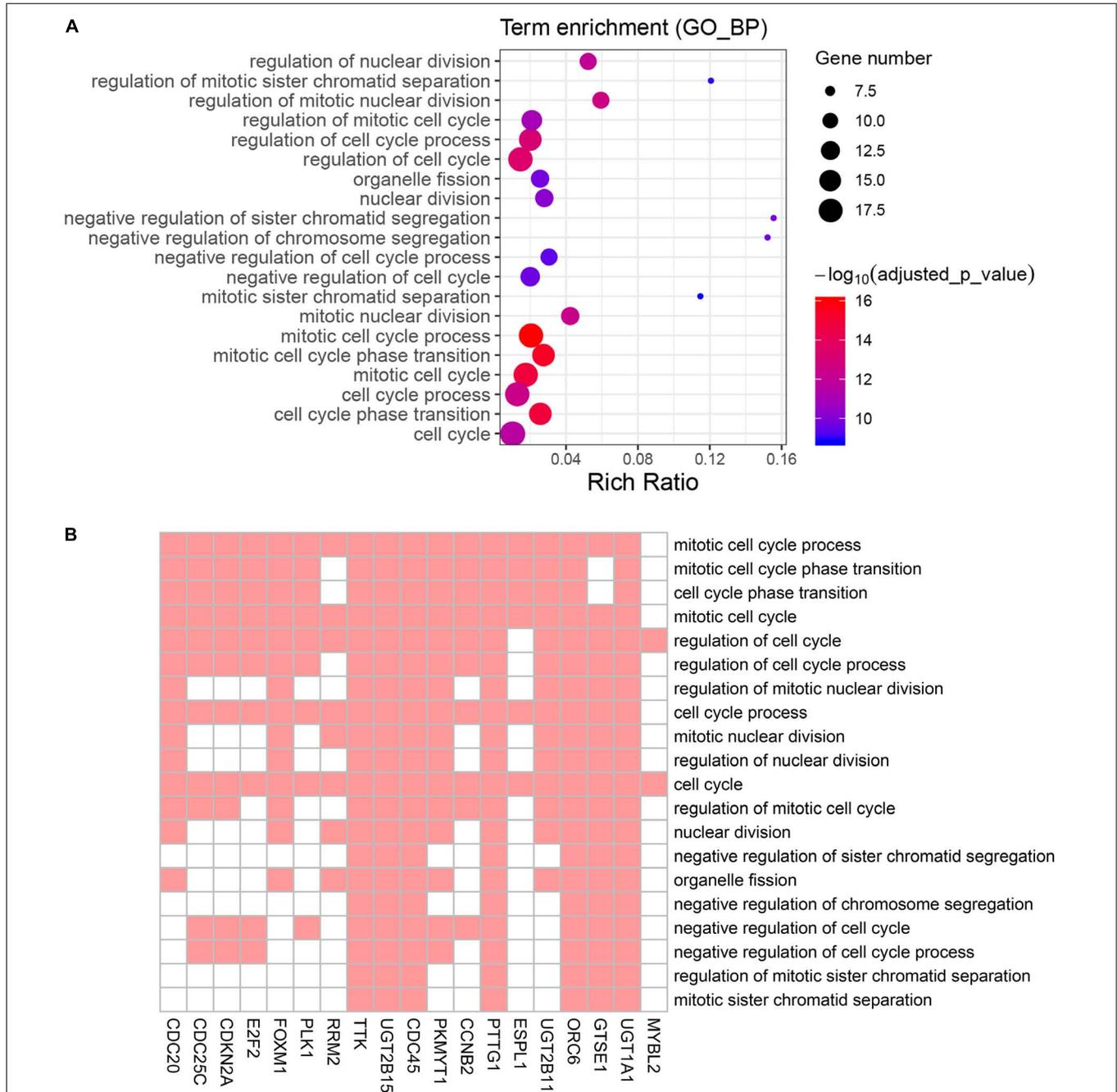
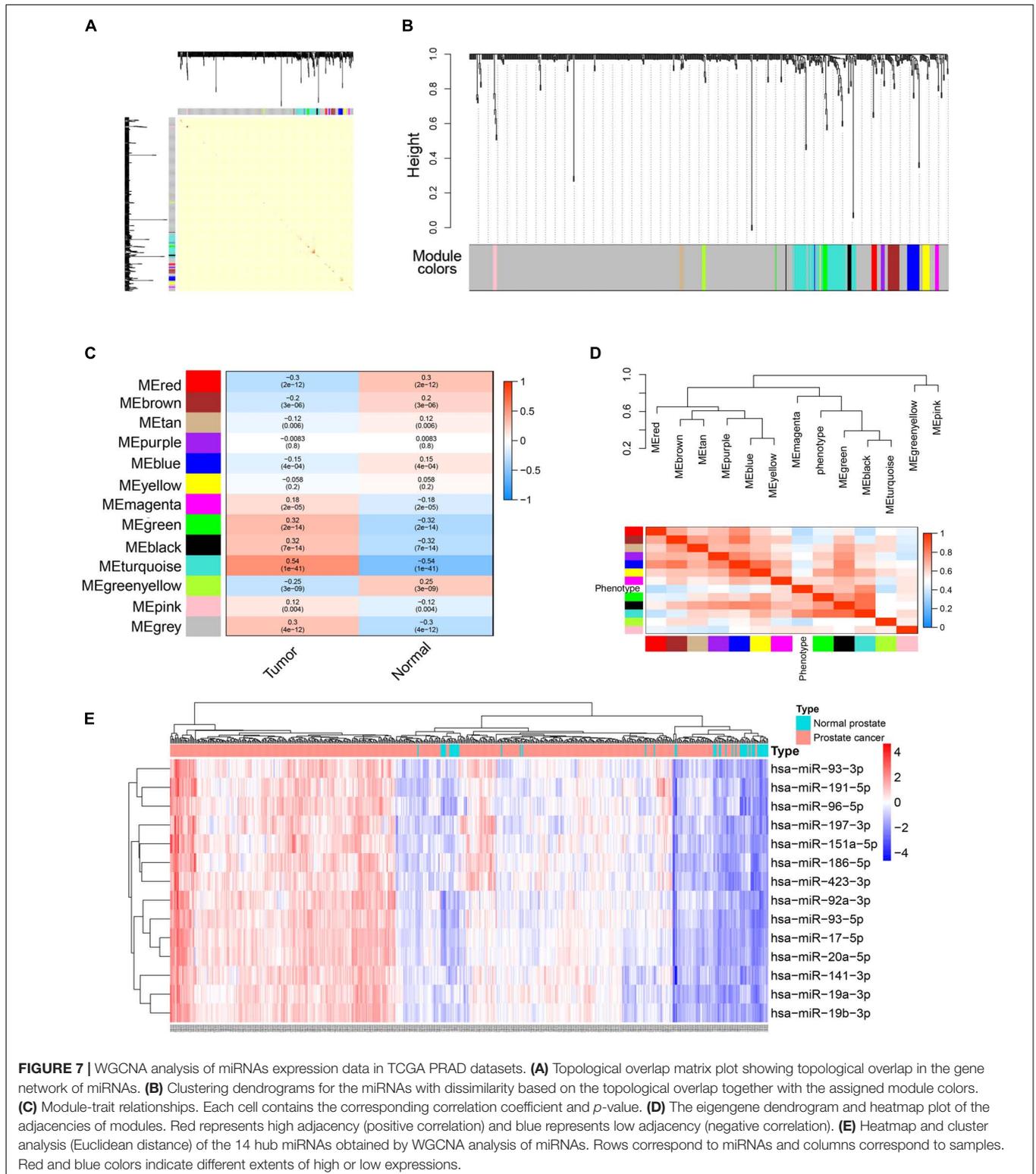


FIGURE 5 | Gene Ontology (GO) biological process enrichment analysis of the genes screened by KEGG pathway analysis. **(A)** Advanced bubble chart shows enrichment of 19 genes in top 20 (based on adjusted *p*-value) enriched GO biological processes. Y-axis labels correspond to biological processes, and X-axis labels correspond to rich ratio (number of input genes in this biological process/number of all genes in this biological process). Color and size of the bubble correspond to enrichment significance and amount of input genes enriched in biological process, respectively. **(B)** Heatmap of the 19 genes enriched in significant GO biological processes. Red boxes mean that the corresponding gene (X-axis) is enriched in the corresponding biological process (Y-axis).

PLK1, FOXM1, CDC25C, PKMYT1, GTSE1, and CDC20 in nine correlated pathways (mitotic G1-G1/S phases, oncogene induced senescence, G2/M transition, mitotic G2-G2/M phases, cell cycle checkpoints, G2/M checkpoints, G1/S-specific transcription,

G1/S transition, and G2/M DNA replication checkpoint) were finally identified and speculated to be potentially correlated with prostate cancer (**Figure 6B**) (Nakayama and Nakayama, 2006; Malumbres and Barbacid, 2009).



WGCNA of miRNAs

To better understand the roles of miRNAs in prostate cancer, we further performed WGCNA analysis for miRNAs based on the expression data obtained from TCGA database. Because the total count of miRNA was far fewer than that of genes, differential expression analysis was not conducted before WGCNA. Similar to the analysis pipeline of WGCNA for genes, a one-step network was constructed (Figures 7A,B) and correlations between trait and modules as well as intramodular correlations were calculated (Figures 7C,D). The turquoise module had the highest correlation coefficient. The connectivity of each miRNA with other miRNAs was calculated and 14 hub miRNAs (miRNAs with the highest module membership) were selected according to the following criteria: miRNA significance value (calculated by WGCNA and representing the correlation between miRNA expression and trait) > 0.3; and absolute value of module membership (calculated by WGCNA and representing the degree each miRNA is associated with the module) > 0.8 (Figure 7E). Based on this analysis, these miRNAs were very likely to interact with target genes and function in the development of prostate cancer.

Construction of Gene–miRNA Interaction Network and Analysis of Hub Genes and miRNAs

To investigate the interactions between the screened miRNAs and genes, we used the miRNet tool which contains miRNA–gene interaction data collected from miRTarBase v7.0, TarBase v7.0, and miRecords and determined the miRNAs that were predicted or validated to target the 13 genes obtained after pathway enrichment analyses. Among all 339 miRNAs, 66 targeted at least two DEGs and were used to find intersections with the 14 hub miRNAs. Five miRNAs (*hsa-mir-17-5p*, *hsa-mir-20a-5p*, *hsa-mir-92a-3p*, *hsa-mir-93-5p*, and *hsa-mir-186-5p*) were commonly shared. Then the interactions between the 5 miRNAs and the aforementioned 13 DEGs were extracted from the miRNA–gene interaction data analyzed using the miRNet tool. According to miRNet, 7 genes (*RRM2*, *ORC6*, *E2F2*, *FOXM1*, *PKMYT1*, *CDC20*, and *MYBL2*) among the 13 DEGs had interactions with the 5 miRNAs. The gene–miRNA interactions of the five miRNAs and seven genes were plotted using Cytoscape software (Figure 8). Three genes (*E2F2*, *RRM2*, and *PKMYT1*) had three or more interactions with miRNAs and four miRNAs (*hsa-mir-17-5p*, *hsa-mir-20a-5p*, *hsa-mir-92a-3p*, and *hsa-mir-93-5p*) had interactions with three genes. Given the high connectivity of the three genes and four miRNAs, it is reasonable to presume that they might be a potential signature in prostate cancer.

Investigation on the Clinical Prognostic Significance of Signature Genes and Expression of Signature miRNAs in Gene Expression Omnibus Datasets

The clinical prognostic significance was studied for *RRM2*, *PKMYT1*, and *E2F2*. According to Kaplan–Meier survival

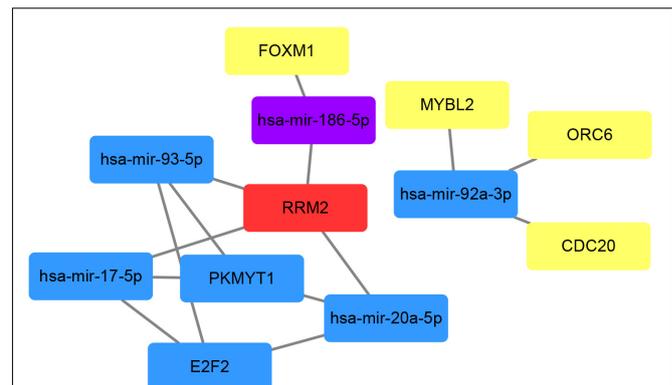
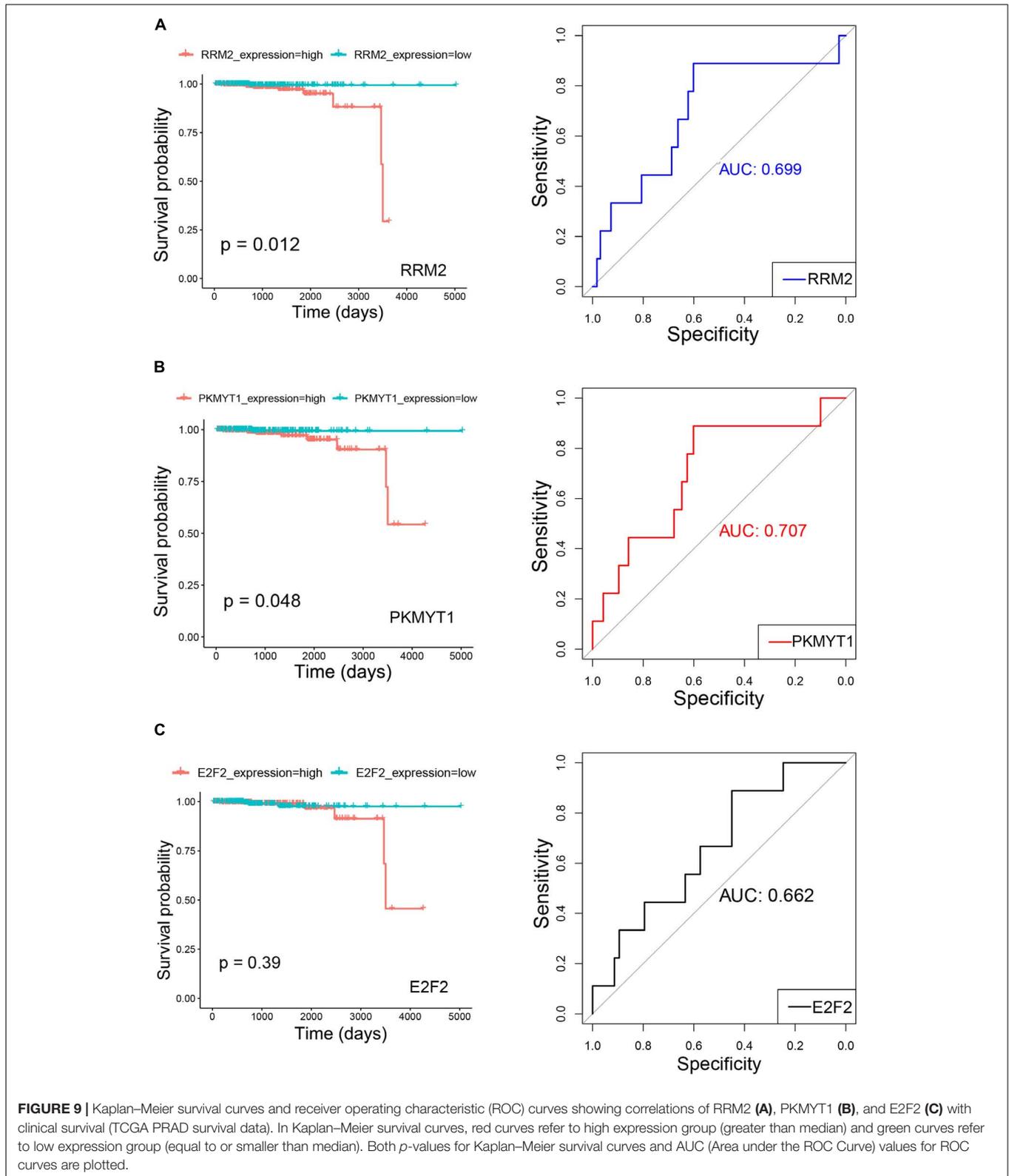


FIGURE 8 | Schematic plot of the gene–miRNA interactions based on miRNet. MiRNet catalogues predicted and validated miRNA–gene interactions. MiRNAs targeting at least two genes were selected to find intersections with the hub miRNAs obtained by WGCNA analysis. Red represents four interactions with other genes or miRNAs in the network, while blue for 3, purple for 2, and yellow for 1.

analyses and survival ROC curve analyses, higher expression of *RRM2* or *PKMYT1* was significantly correlated with worse survival ($p < 0.05$, $AUC > 0.69$) (Figures 9A,B). For *E2F2*, higher expression did not indicate worse survival but ROC curve still showed considerable AUC value (Figure 9C). To conduct some validation of sorts, we further studied the expression of four signature miRNAs in several Gene Expression Omnibus (GEO) datasets. The GEO database is an independent public repository that archives and freely distributes high-throughput gene expression and other functional genomics datasets (Barrett et al., 2013). All four miRNAs were more highly expressed in prostate cancer than in normal prostate tissues and this was confirmed by at least three GEO datasets for each miRNA (Figure 10). Among the four miRNAs, three members (*hsa-mir-17-5p*, *hsa-mir-20a-5p*, and *hsa-mir-93-5p*) had some of the strongest p -values for tumor and normal differential expression in the independent GEO datasets (GSE21036, and GSE36802). All of these data provide strong evidence that the signatures of three genes and four miRNAs play crucial roles in the development of prostate cancer.

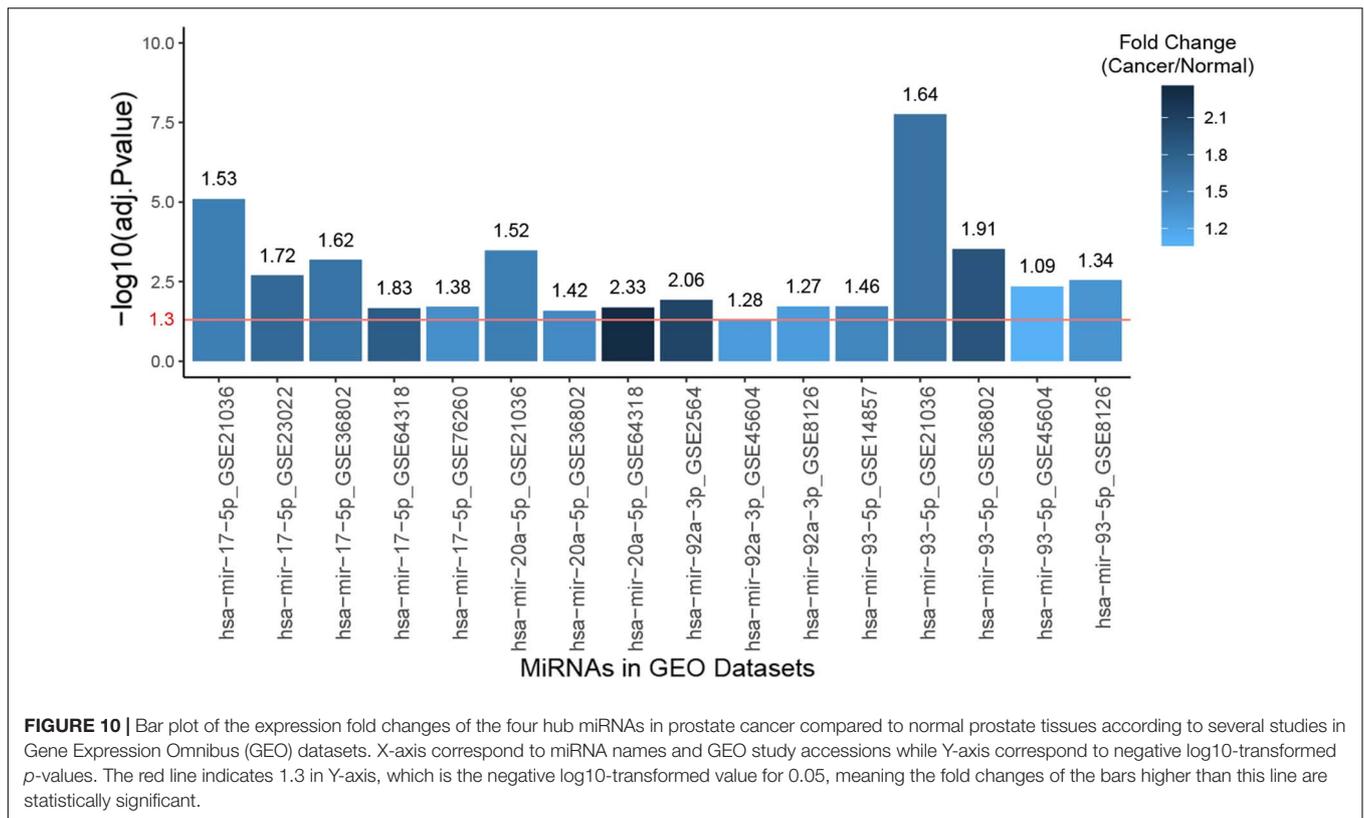
DISCUSSION

With the increasing number of studies focusing on prostate cancer, much progress has been made in the treatment and understanding of the mechanisms underlying this disease. However, controversies remain about several issues such as the screening for prostate cancer and the treatment of localized disease (Pignot et al., 2018). The optimum management of prostate cancer is challenging partly due to the heterogeneity and sometimes indolent nature of the disease. Detailed biology of oncogenesis and progression of prostate cancer, especially advanced subtypes, still needs to be elucidated for screening, stratification and treatment strategies. Thus, a signature with prognostic and therapeutic implications is of great significance.



Integrative analysis of accumulated expression data can be performed to obtain more reliable information and more feasible measures for determining the potential diagnostic

and prognostic signatures of tumors and to investigate the molecular mechanisms (Robertson et al., 2017; Li et al., 2018; Ye et al., 2018). WGCNA builds network modeling that relies



on statistical methods and improves simple correlation networks by quantifying not only the correlations between individual pairs of genes, but also the extent to which these genes share the same neighbors (Zhang and Horvath, 2005; DiLeo et al., 2011). WGCNA provides an efficient approach through which the effects of phenotype can be detected in modeled networks (Langfelder and Horvath, 2008). Many studies have indicated that cancer biology is regulated by miRNAs. Research works focusing on miRNA-based cancer treatments are gaining increasing attention due to the ability of miRNAs to concurrently target multiple effectors of pathways involved in cell differentiation, proliferation, and survival (Garzon et al., 2010). In this study, using the expression data of genes and miRNAs of prostate cancer in TCGA database as well as the WGCNA approach, we identified three genes and four miRNAs that were tightly correlated with prostate cancer and clinical prognosis.

In our pathway enrichment analyses, we utilized three databases, namely KEGG, GO, and REACTOME. For each kind of functional enrichment analysis, there were always several pathways relevant to the cell cycle that were enriched. It is universally accepted that dysregulation of the cell cycle is characteristic of tumors, which further proved that our functional enrichment analyses were properly conducted (Malumbres and Barbacid, 2009).

Protein kinase membrane-associated tyrosine/threonine 1 (*PKMYT1*) inhibits progression of the cell cycle through the phosphorylation of cyclin-dependent kinase 1 (*CDK1*) (Choi et al., 2009). It is well-established that *CDK1* is a key molecule

in the progression of mitosis, and thus *PKMYT1* is thought to play important roles in the regulation of the cell cycle and tumor biology (Parker and Piwnicka-Worms, 1992). *PKMYT1* is a crucial promoter in the development of hepatocellular carcinoma (Aguilar et al., 2017; Liu et al., 2017), and overexpression of *PKMYT1* indicates a poor prognosis and enhances proliferation and tumorigenesis in non-small cell lung cancer (Sun et al., 2019). Currently the functional significance of *PKMYT1* in prostate cancer remains unclear, and it makes sense to determine whether *PKMYT1* is indispensable in prostate cancer, which requires further investigation. The ribonucleotide reductase M2 subunit (*RRM2*) plays important role in several human cancers, including colorectal cancer, hepatocellular carcinoma, pancreatic adenocarcinoma, and breast cancer (Lee et al., 2014; Sierzega et al., 2017; Nana et al., 2018; Chen et al., 2019). Increased expression of *RRM2* has been demonstrated to be a mechanism driving poor patient outcomes in prostate cancer, in accordance with our findings (Mazzu et al., 2019). According to our study, the high expression of *PKMYT1* and *RRM2* was correlated with prostate cancer oncogenesis and poor prognosis, indicating their potential oncogenic roles in prostate cancer. *E2F* transcription factor 2 (*E2F2*) is one of the activators of the family that controls the transition between the G1 and S phase through various upstream signals (Sherr, 1996), and functions as an oncogene in lung cancer (Feliciano et al., 2017). Li et al. (2017), reported that *E2F2* variants are predictive biomarkers for recurrence risk in patients with squamous cell carcinoma of the oropharynx. It is documented that downregulated *E2F2*

is correlated with inhibition of cell proliferation in prostate cancer (Dong et al., 2010). However, more research regarding its roles in prostate cancer is still needed. Among the four hub miRNAs found in this study, *miRNA-17-5p* has been extensively investigated (Hussein et al., 2009; Dellago et al., 2017). *MiRNA-17-5p* is shown by multiple studies to be at the crossroads of aging and cancer and might be a promising biomarker or even therapeutic tool and target in tumors (Dellago et al., 2017). The high expression of *miRNA-17-5p* is significantly associated with the progression of various cancers probably through increasing the expression of *p53* (Thompson et al., 2016; Duan et al., 2018). However, *miRNA-17-5p* also reportedly acts as a tumor suppressor in triple-negative breast cancer, which is the breast cancer subtype with the poorest prognosis (Wang et al., 2019). Aberrant expression of *miRNA-92a* has been reported in various tumors and as such has potential value as a tumor marker or novel target for cancer treatment (Li et al., 2014). In prostate cancer, the expression of *miR-92a* is markedly overexpressed according to a large-scale miRnome analysis (Volinia et al., 2006). For each gene or miRNA included in the signature found in this study, our report provided novel targets, or at least presented further confirmation of previously published findings of prostate cancer.

It is worth noting that all exploratory data analysis approaches such as WGCNA require further validation to confirm the putative molecular networks. In addition, the expression of a specific gene or miRNA might vary among different subtypes of prostate cancer, especially taking into account the heterogeneous nature of this disease (Boyd et al., 2012).

A signature with prognostic significance makes great sense for prostate cancer patients. For instance, a signature identifying prostate cancer patients at risk of dying and

who can be safely observed is of tremendous importance, as we are aware of overtreatment concerns and more patients are being placed on active surveillance protocols. This study preliminarily investigated potential candidate drivers in prostate cancer as well as the underlying mRNA-miRNA interactions, providing potential therapeutic targets for prostate cancer. It will be of great interest to determine whether this signature will prove useful and gain traction in further studies.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary files.

AUTHOR CONTRIBUTIONS

JW, YY, QD, JZ, YW, GY, JY, and YT: conception. JW, YY, QD, JZ, and YW: data collection and analyses. JW, YY, and QD: writing, review, and/or revision of the manuscript. GY, JY, and YT: study supervision.

FUNDING

This work was supported by the National Natural Science Foundation of China (YT, #81571432), Hunan Natural Science Foundation (YT, #2019JJ80097 and YW, #2019JJ80008), and Research Project of Hunan Health and Family Planning Commission (GY, #20180895).

REFERENCES

- Aguiar, P. N. Jr., De Mello, R. A., Hall, P., Tadokoro, H., and Lima Lopes, G. (2017). PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: updated survival data. *Immunotherapy* 9, 499–506. doi: 10.2217/imt-2016-0150
- Attard, G., Parker, C., Eeles, R. A., Schroder, F., Tomlins, S. A., Tannock, I., et al. (2016). Prostate cancer. *Lancet* 387, 70–82.
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995.
- Boyd, L. K., Mao, X., and Lu, Y. J. (2012). The complexity of prostate cancer: genomic alterations and heterogeneity. *Nat. Rev. Urol.* 9, 652–664. doi: 10.1038/nrurol.2012.185
- Calcinotto, A., Kohli, J., Zagato, E., Pellegrini, L., Demaria, M., and Alimonti, A. (2019). Cellular senescence: aging, cancer, and injury. *Physiol. Rev.* 99, 1047–1078. doi: 10.1152/physrev.00020.2018
- Chen, W. X., Yang, L. G., Xu, L. Y., Cheng, L., Qian, Q., Sun, L., et al. (2019). Bioinformatics analysis revealing prognostic significance of RRM2 gene in breast cancer. *Biosci. Rep.* 30:39.
- Choi, H. S., Bode, A. M., Shim, J. H., Lee, S. Y., and Dong, Z. (2009). c-Jun N-terminal kinase 1 phosphorylates Myt1 to prevent UVA-induced skin cancer. *Mol. Cell. Biol.* 29, 2168–2180. doi: 10.1128/mcb.01508-08
- Chou, C. H., Shrestha, S., Yang, C. D., Chang, N. W., Lin, Y. L., Liao, K. W., et al. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 46, D296–D302.
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44:e71. doi: 10.1093/nar/gkv1507
- Culig, Z., and Santer, F. R. (2014). Androgen receptor signaling in prostate cancer. *Cancer Metastasis Rev.* 33, 413–427.
- Dellago, H., Bobbili, M. R., and Grillari, J. (2017). MicroRNA-17-5p: at the crossroads of cancer and aging – A mini-review. *Gerontology* 63, 20–28. doi: 10.1159/000447773
- DiLeo, M. V., Strahan, G. D., den Bakker, M., and Hoekenga, O. A. (2011). Weighted correlation network analysis (WGCNA) applied to the tomato fruit metabolome. *PLoS ONE* 6:e26683. doi: 10.1371/journal.pone.0026683
- Dong, Q., Meng, P., Wang, T., Qin, W., Qin, W., Wang, F., et al. (2010). MicroRNA let-7a inhibits proliferation of human prostate cancer cells in vitro and in vivo by targeting E2F2 and CCND2. *PLoS ONE* 5:e10147. doi: 10.1371/journal.pone.0010147
- Duan, F., Yang, Y., Liu, W., Zhao, J., Song, X., Li, L., et al. (2018). Quantifying the prognostic significance of microRNA-17/17-5P in cancers: a meta-analysis based on published studies. *Cancer Manage. Res.* 10, 2055–2069. doi: 10.2147/cmar.s163415
- Duell, E. J., Lujan-Barroso, L., Sala, N., Deitz, McElyea, S., Overvad, K., et al. (2017). Plasma microRNAs as biomarkers of pancreatic cancer risk in a prospective cohort study. *Int. J. Cancer* 141, 905–915.
- Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., et al. (2018). The reactome pathway knowledgebase. *Nucleic Acids Res.* 46, D649–D655.

- Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnaud, V., et al. (2017). Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics* 18:142. doi: 10.1186/s12859-017-1559-2
- Feliciano, A., Garcia-Mayea, Y., Jubierre, L., Mir, C., Hummel, M., Castellvi, J., et al. (2017). miR-99a reveals two novel oncogenic proteins E2F2 and EMR2 and represses stemness in lung cancer. *Cell Death Dis.* 8:e3141. doi: 10.1038/cddis.2017.544
- Garzon, R., Marcucci, G., and Croce, C. M. (2010). Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat. Rev. Drug Discov.* 9, 775–789. doi: 10.1038/nrd3179
- Han, D., Li, J., Wang, H., Su, X., Hou, J., Gu, Y., et al. (2017). Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression. *Hepatology* 66, 1151–1164. doi: 10.1002/hep.29270
- Hussein, K., Dralle, W., Theophile, K., Kreipe, H., and Bock, O. (2009). Megakaryocytic expression of miRNA 10a, 17-5p, 20a and 126 in Philadelphia chromosome-negative myeloproliferative neoplasm. *Ann. Hematol.* 88, 325–332. doi: 10.1007/s00277-008-0602-9
- Inui, M., Martello, G., and Piccolo, S. (2010). MicroRNA control of signal transduction. *Nat. Rev. Mol. Cell Biol.* 11, 252–263.
- Kastan, M. B., and Bartek, J. (2004). Cell-cycle checkpoints and cancer. *Nature* 432, 316–323.
- Lange, C. A., Gioeli, D., Hammes, S. R., and Marker, P. C. (2007). Integration of rapid signaling events with steroid hormone receptor action in breast and prostate cancer. *Annu. Rev. Physiol.* 69, 171–199. doi: 10.1146/annurev.physiol.69.031905.160319
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Lee, B., Ha, S. Y., Song, D. H., Lee, H. W., Cho, S. Y., and Park, C. K. (2014). High expression of ribonucleotide reductase subunit M2 correlates with poor prognosis of hepatocellular carcinoma. *Gut Liver* 8, 662–668. doi: 10.5009/gnl13392
- Li, D., Hao, X., and Song, Y. (2018). Identification of the key MicroRNAs and the miRNA-mRNA regulatory pathways in prostate cancer by bioinformatics methods. *BioMed Res. Int.* 2018:6204128.
- Li, M., Guan, X., Sun, Y., Mi, J., Shu, X., Liu, F., et al. (2014). miR-92a family and their target genes in tumorigenesis and metastasis. *Exp. Cell Res.* 323, 1–6. doi: 10.1016/j.yexcr.2013.12.025
- Li, Y., Sturgis, E. M., Zhu, L., Cao, X., Wei, Q., Zhang, H., et al. (2017). E2F transcription factor 2 variants as predictive biomarkers for recurrence risk in patients with squamous cell carcinoma of the oropharynx. *Mol. Carcinogenesis* 56, 1335–1343. doi: 10.1002/mc.22595
- Litwin, M. S., and Tan, H. J. (2017). The diagnosis and treatment of prostate cancer: a review. *JAMA* 317, 2532–2542.
- Liu, J., Li, Y., Lu, Z., Gu, J., Liang, Y., Huang, E., et al. (2019). Deceleration of glycometabolism impedes IgG-producing B-cell-mediated tumor elimination by targeting SATB1. *Immunology* 156, 56–68. doi: 10.1111/imm.12998
- Liu, L., Wu, J., Wang, S., Luo, X., Du, Y., Huang, D., et al. (2017). PKMYT1 promoted the growth and motility of hepatocellular carcinoma cells by activating beta-catenin/TCF signaling. *Exp. Cell Res.* 358, 209–216. doi: 10.1016/j.yexcr.2017.06.014
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.
- Malumbres, M., and Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* 9, 153–166. doi: 10.1038/nrc2602
- Mazzu, Y. Z., Armenia, J., Chakraborty, G., Yoshikawa, Y., Coggins, S. A., Nandakumar, S., et al. (2019). A novel mechanism driving poor-prognosis prostate cancer: overexpression of the DNA repair gene, ribonucleotide reductase small subunit M2 (RRM2). *Clin. Cancer Res.* 25, 4480–4492. doi: 10.1158/1078-0432.ccr-18-4046
- Munoz-Fontela, C., Mandinova, A., Aaronson, S. A., and Lee, S. W. (2016). Emerging roles of p53 and other tumour-suppressor genes in immune regulation. *Nat. Rev. Immunol.* 16, 741–750. doi: 10.1038/nri.2016.99
- Nakayama, K. I., and Nakayama, K. (2006). Ubiquitin ligases: cell-cycle control and cancer. *Nat. Rev. Cancer* 6, 369–381. doi: 10.1038/nrc1881
- Nana, A. W., Wu, S. Y., Yang, Y. S., Chin, Y. T., Cheng, T. M., Ho, Y., et al. (2018). Nano-diamino-tetrac (NDAT) enhances resveratrol-induced antiproliferation by action on the RRM2 pathway in colorectal cancers. *Hormones Cancer* 9, 349–360. doi: 10.1007/s12672-018-0334-9
- Parker, L. L., and Piwnica-Worms, H. (1992). Inactivation of the p34cdc2-cyclin B complex by the human WEE1 tyrosine kinase. *Science* 257, 1955–1957. doi: 10.1126/science.1384126
- Pignot, G., Mailliet, D., Gross, E., Barthelemy, P., Beauval, J. B., Constans-Schlurmann, F., et al. (2018). Systemic treatments for high-risk localized prostate cancer. *Nat. Rev. Urol.* 15, 498–510.
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47, W191–W198.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gk_v007
- Robertson, A. G., Shih, J., Yau, C., Gibb, E. A., Oba, J., Mungall, K. L., et al. (2017). Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer Cell* 32:e215.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77
- Rupaimoole, R., and Slack, F. J. (2017). MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat. Rev. Drug Discov.* 16, 203–222. doi: 10.1038/nrd.2016.246
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Sherr, C. J. (1996). Cancer cell cycles. *Science*. 274, 1672–1677. doi: 10.1126/science.274.5293.1672
- Sierzega, M., Pach, R., Kulig, P., Legutko, J., and Kulig, J. (2017). Prognostic implications of expression profiling for gemcitabine-related genes (hENT1, dCK, RRM1, RRM2) in patients with resectable pancreatic adenocarcinoma receiving adjuvant chemotherapy. *Pancreas* 46, 684–689. doi: 10.1097/MPA.0000000000000807
- Sun, Q. S., Luo, M., Zhao, H. M., and Sun, H. (2019). Overexpression of PKMYT1 indicates the poor prognosis and enhances proliferation and tumorigenesis in non-small cell lung cancer via activation of Notch signal pathway. *Eur. Rev. Med. Pharmacol. Sci.* 23, 4210–4219.
- The TCGA Legacy, (2018). The TCGA Legacy. *Cell* 173, 281–282. doi: 10.1016/j.cell.2018.03.049
- Thompson, M. A., Edmonds, M. D., Liang, S., McClintock-Treep, S., Wang, X., Li, S., et al. (2016). miR-31 and miR-17-5p levels change during transformation of follicular lymphoma. *Hum. Pathol.* 50, 118–126. doi: 10.1016/j.humpath.2015.11.011
- Vlachos, I. S., Paraskevopoulou, M. D., Karagkouni, D., Georgakilas, G., Vergoulis, T., Kanellos, I., et al. (2015). DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res.* 43, D153–D159.
- Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., et al. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2257–2261. doi: 10.1073/pnas.0510565103
- Vorvis, C., Koutsoumpa, M., and Iliopoulos, D. (2016). Developments in miRNA gene signaling pathways in pancreatic cancer. *Future Oncol.* 12, 1135–1150. doi: 10.2217/fo-2015-0050
- Wang, S., Oh, D. Y., Leventaki, V., Drakos, E., Zhang, R., Sahin, A. A., et al. (2019). MicroRNA-17 acts as a tumor chemosensitizer by targeting JAB1/CSN5 in triple-negative breast cancer. *Cancer Lett.* 465, 12–23. doi: 10.1016/j.canlet.2019.08.016
- Winbanks, C. E., Ooi, J. Y., Nguyen, S. S., McMullen, J. R., and Bernardo, B. C. (2014). MicroRNAs differentially regulated in cardiac and skeletal muscle in health and disease: potential drug targets? *Clin. Exp. Pharmacol. Physiol.* 41, 727–737.
- Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X., and Li, T. (2009). miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.* 37, D105–D110.

- Xu, T., Su, N., Liu, L., Zhang, J., Wang, H., Zhang, W., et al. (2018). miRBaseConverter: an R/Bioconductor package for converting and retrieving miRNA name, accession, sequence and family information in different versions of miRBase. *BMC Bioinformatics* 19:514. doi: 10.1186/s12859-018-2531-5
- Yang, Z., Wu, L., Wang, A., Tang, W., Zhao, Y., Zhao, H., et al. (2017). dbDEMC 2.0: updated database of differentially expressed miRNAs in human cancers. *Nucleic Acids Res.* 45, D812–D818.
- Ye, Y., Li, S. L., and Wang, S. Y. (2018). Construction and analysis of mRNA, miRNA, lncRNA, and TF regulatory networks reveal the key genes associated with prostate cancer. *PLoS ONE* 13:e0198055. doi: 10.1371/journal.pone.0198055
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic*s 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statist. Appl. Genet. Mol. Biol.* 4:17.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wei, Yin, Deng, Zhou, Wang, Yin, Yang and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.