Introduction

Prostate Cancer is a high incidence malignant tumor that affects men worldwide. In China, prostate cancer accounted for 8.16% of the total number of new cancer cases in males, with a mortality rate of 13.61% (1). In the United States, prostate cancer is the most common (comprising 26% of all cases) malignant tumor in males (2). The incidence of prostate cancer has continuously increased globally (3). Prostate cancer patients undergo biochemical relapse (BCR), which results in the poor prognosis (4,5). BCR is defined as two consecutive rising prostate-specific antigen (PSA) values >2 ng/mL above the nadir PSA value after radiation therapy (RT) or >0.2 ng/mL after radical prostatectomy (RP) (6).

Urologists have made great progress in the diagnosis and treatment of prostate cancer, especially with the...
discovery of new biomarkers for liquid biopsies and imaging techniques that can assist clinicians in making decisions. Prostate cancer is a group of heterogeneous tumors with individual differences, molecular markers play an important role in assisting diagnosis and predicting prognosis (7,8). Therefore, it is essential to find novel and effective molecular markers associated with the diagnosis and prognosis for PCa.

Non-coding microRNAs (miRNAs) are single-stranded small RNA molecules involved in post-transcriptional gene regulation (9). Accumulating evidence has revealed that miRNAs participate in regulating cellular biological and pathophysiological processes of cancer cells. In addition, miRNAs can be detected stably in biofluids to achieve non-invasive tumor diagnosis and efficacy evaluation, which have been applied to assist clinical diagnosis and prognosis. However, few studies on multi-dataset analysis of combined RRA of miRNA-mRNA in prostate cancer have been reported. This study aims to construct miRNA-mRNA regulatory networks of prostate cancer progression by using bioinformatics methods, to identify core miRNAs and hub target genes. Hub genes may be used to build-up a nomogram model for clinical prognosis. We present the following article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-653/rc).

**Methods**

**Data collection and preprocessing**

The datasets were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). We selected 3 miRNA datasets (GSE60117, GSE21036 and GSE76260), 4 mRNA datasets (GSE79021, GSE62872, GSE70768 and GSE70769), and The Cancer Genome Atlas (TCGA) database (https://gdc-portal.nci.nih.gov/). The characteristics of these datasets were shown in Table 1. PCa patients were confirmed by pathological examination and included both metastatic and *in situ* cancers. In GSE70769 and TCGA datasets, the survival information of the patient who had undergone RP or RT was collected for the subsequent biochemical relapse-free survival (BCRFS) analysis. The study was conducted based on the Declaration of Helsinki (as revised in 2013).

**Differential miRNA and mRNA screening**

Differentially expressed miRNAs (DEMs) and differentially expressed genes (DEGs) were determined by comparing two groups between normal and cancerous tissues in each GEO dataset. Datasets were corrected using the ‘normalizeBetweenArrays’ package of R software (version 4.1.2). The ‘limma’ package of R software was used to screen DEGs in the dataset, and the screening criteria for DEMs and DEGs were adjusted P values <0.05. The correction methods were Benjamini and Hochberg (BH), to address false-positivity (10). The DEMs and DEGs in these datasets were analyzed using the R package ‘RobustRankAggreg’ (11). The RRA analysis was applied to calculate the significance probability of all genes in the final sequence, which might integrate the differential expression matrix of each dataset (11).
Construction of miRNA-mRNA regulatory networks

FunRich was a bioinformatics tool, which could predict miRNA target genes for functional enrichment (12). Common DEGs (co-DEGs) could be obtained by using FunRich (version 3.1.3) to predict miRNA target genes. Co-DEGs were used to construct miRNA-mRNA regulatory networks. Cytoscape (version 3.8.2) was used to visually display each network (13). Top 3 miRNAs with Edge Percolated Component (EPC) scores (connectivity greater than 30) were selected to build-up the core network of miRNAs using the EPC topology algorithm. Co-DEGs regulated by miRNAs in the core network were selected as core network genes.

Gene function enrichment analysis

The GO (Gene Ontology) function and KEGG (Kyoto Encyclopedia of Genes and Genomes) signal pathway analysis of co-DEGs were performed using the R software package “clusterProfiler” (14). GO enrichment included biological process (BP), cellular component (CC) and molecular function (MF). For KEGG functional enrichment-related signalling pathways, the screening criteria were adjusted P<0.05 to control false discovery rate (FDR).

Differential expression and prognostic value of hub regulatory genes

TCGA and Genotype-Tissue Expression (GTEx) databases contain clinical samples and demographic/clinical information. Therefore, they can be used to verify statistical differences in the expression levels of core network genes between cancer and normal tissues in a large cohort. Hub target genes were analyzed using GEPIA2 web tools (http://gepia2.cancer-pku.cn/).

Establishment of nomogram

The risk score of patients was calculated by ridge regression analysis. The TCGA and GSE70769 cohort was treated as training and validation set, respectively. Multivariate and univariate COX regression analyses were performed for each candidate. A nomogram prediction model was established for independent prognostic factors using R package ‘rms’ (https://CRAN.R-project.org/package=rms). 100 points were allocated to the most dangerous predictors with the highest B coefficient, and corresponding scores were assigned to other predictors according to relative weight. The accuracy of prediction was evaluated using calibration diagram and AUC.

Statistical analysis

Statistical analysis was performed using R software (version 4.1.2). Differences between groups were compared with Wilcoxon tests. The BCRFS curves were generated by using Kaplan-Meier (K-M) plots, while the differences were evaluated by log-rank tests. Univariate and multivariate Cox proportional hazard models were applied to estimate potentially predictive values of parameters. Time-ROC curves were plotted and utilized to assess potential prognostic power of risk scores. Calibration plots and AUCs were applied to evaluate the prediction accuracy of a nomogram. A P<0.05 was considered statistically significant.

Results

Screening of differential miRNAs and mRNAs associated with prostate cancer

Three miRNA datasets were analyzed with the RRA method. A total of 22 DEMs were up-regulated whereas 14 down-regulated (Figure 1A). A total of 896 DEGs were up-regulated whereas 915 down-regulated in prostate cancer. As shown in Figure 1B, the top 20 up-regulated and down-regulated genes were sorted by Log2 Fold Change.

Construction of the miRNA-mRNA network

Among the 36 DEMs screened from miRNA datasets, 17 miRNAs could be predicted by FunRich software, and a total of 2,575 target genes could be predicted. Among them, 184 genes were selected based on intersection of co-DEGs and target genes. Then, miRNA-mRNA reaction network was constructed according to regulatory relationship using Cytoscape3.8.2 (Figure 2A). In addition, three core network miRNAs were obtained by EPC topology algorithm screening, which were highly expressed in prostate cancer tissues, including hsa-miR-106b-5p, hsa-miR-17-5p and hsa-miR-183-5p. The key gene network comprised of 32 core genes as shown in Cytoscape (Figure 2B), while the key network module information was summarized in Table 2.
GO and KEGG enrichment analysis

GO analysis of genes included biological process (BP), molecular function (MF) and cellular component (CC) (Figure 3A). The biological processes of co-DEGs mainly enriched in actin filaments regulation, mesenchymal development and muscular system. In terms of cellular components, “focal adhesions”, “cell-substrate adherens junction”, and “cell-substrate junctions” were the most dominant elements. For molecular functions, the GO terms enriched with co-DEGs included actin binding, platelet-derived growth factor receptor binding, and activation of 3’5’ cyclic-AMP phosphodiesterase. In addition, KEGG analysis indicated involvement of focal adhesion, cGMP-PKG signaling pathway and actin skeleton regulatory pathway in the pathogenesis of prostate cancer (Figure 3B). Afterwards, co-DEGs were mapped to 18 KEGG pathways, as shown in Figure 3C.

Validation of hub target genes

We performed gene expression values analyses of the four hub target genes (TMEM100, FRMD6, NBL1 and STARD4) by GEPIA2. These gene expression levels in prostate cancer were lower than those in normal control (P<0.05), as shown in Figure 4.

Establishment and verification of the nomogram

A risk score was calculated by ridge regression. The formula was as follows: risk score = (-0.018372 × STARD4) + (-0.017621 × FRMD6) + (-0.000546 × NBL1) + (-0.000064 × TMEM100). According to univariate and multivariate COX regression analyses, Gleason and risk scores might be used as independent prognostic factors (Figure 5A). Nomogram can predict BCRFS for PCa patients (Figure 5B). The calibration plot demonstrated
good consistency between the prediction by nomogram and actual observation of 1-, 3- and 5-year BCRFS in prostate cancer (Figure 5C). The novel nomogram predicted the AUC of BCRFS at 1-, 3- and 5-year of 0.713, 0.732 and 0.753 (Figure 5D). The BCRFS time of high-risk group was significantly lower than that of the low-risk group (Figure 5E).

**Discussion**

Previous studies have indicated that aberrant expression of miRNAs may dysregulate expression of tumour-related genes, and thus affect tumour development and progression (9,15-17). Although many genes have been proposed as miRNA regulatory targets in PCa, their functions remain unknown. The pathogenesis of prostate cancer involves various signalling pathways regulated by miRNAs-mRNAs. Several miRNAs have been well studied in prostate cancer. For example, circRNA-HIPK3 may upregulate CDC25B and CDC2, promote G2/M transformation and induce cell proliferation of prostate cancer by inhibiting miR-338-3p (18). In addition, overexpression of miR-338-3p may downregulate BCL2 in LNCaP cells, and suppress apoptosis of prostate cancer cells (17). By contrast, MiR-338-3p directly targets the RAB23 gene and inhibits tumour growth in prostate cancer (19). Thus, it is critical to explore inter-regulatory relationship of miRNAs-mRNAs in tumour development and progression of prostate cancer.

In this study, hsa-miR-106b-5p, hsa-miR-17-5p and hsa-miR-183-5p were identified to construct a miRNA-mRNA regulatory network of prostate cancer. Hsa-miR-106b-5p, as an oncogene, may down-regulate RBMS1 expression in prostate cancer tissues and affect cell proliferation, gap closure and colony formation (20). Meanwhile, hsa-miR-106b in the blood of prostate cancer patients is significantly higher than that of healthy controls (21). Moreover, lncRNA-FER1L4 competitively binds a specific miRNA through its miRNA response element and represses the inhibitory effects of hsa-miR-106b-5p on its target mRNAs. This process participates in the post-transcriptional regulation of cancer-related genes (22). In addition, high expression of hsa-miR-17-5p in the epithelium was a
predictive marker for poor prognosis in patients with PCa (23). Furthermore, circRNA-ITCH inhibited PCa progression by sponging hsa-miR-17-5p and increasing HOXB13 expression (24). In addition, high expression of hsa-miR-183-5p in local tissues of advanced prostate cancer may be related to lymph node metastasis and diffusion (25). Collectively, the three core miRNAs in this study were abnormally highly expressed in prostate cancer tissues, which was consistent with previous studies (16,21,23).

Four hub mRNAs were also identified associated with prostate cancer in this study, including TMEM100, FRMD6, NBL1 and STARD4. Previous studies have investigated the relationship between these hub genes and prostate cancer. For example, FRMD6 was abnormally hypermethylated and significantly down-regulated in prostate cancer tissues. Relatively low expression of FRMD6 was associated with postoperative biochemical recurrence. FRMD6 is proposed as a tumour-suppressive gene in prostate cancer (26). In addition, NBL1, a secreted protein, was highly expressed in the normal prostate. Low expression of NBL1 was associated with prostate cancer progression as a candidate biomarker (27). Notably, NBL1 expression in prostate cancer gland was down-regulated compared with normal gland, which might act as a tumour suppressor gene and a new therapeutic target (28). In this study, FRMD6 and NBL1 were significantly downregulated in prostate cancer tissues, and associated with disease-free survival (DFS) in prostate cancer patients. Its regulatory effects on TMEM100 and STARD4 in PCa necessitate further investigations.

Through GO functional enrichment analysis, co-DEGs were significantly enriched in actin filament regulation process, adhesion spots and actin-binding function. KEGG pathway analysis identified co-DEGs enriched signalling pathways, including focal adhesion, cGMP-PKG, actin skeleton regulation, prostate cancer, and PI3K-Akt and cAMP. These cellular functions and signalling pathways were related to tumour development and progression of prostate cancer, which were consistent with previous studies (15,29,30). Meanwhile, inflammatory signalling pathways (PI3K-Akt and cAMP) indicate the importance of inflammation in tumour progression of localized prostate cancer. Inflammation markers of prostate cancer will help to discover accurate biomarkers and effective therapeutic targets (31). More importantly, in the validation set with expanded sample size by using the exogenous database “GEPIA2”, the hub target genes were significantly differentially expressed in prostate cancer versus normal tissues.

### Table 2 miRNA-mRNA core network module information

<table>
<thead>
<tr>
<th>Gene</th>
<th>Category</th>
<th>Change trend</th>
<th>Degree value</th>
<th>EPC score</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-17-5p</td>
<td>miRNA</td>
<td>Up</td>
<td>35</td>
<td>42.675</td>
</tr>
<tr>
<td>hsa-miR-106b-5p</td>
<td>miRNA</td>
<td>Up</td>
<td>34</td>
<td>42.662</td>
</tr>
<tr>
<td>hsa-miR-183-5p</td>
<td>miRNA</td>
<td>Up</td>
<td>44</td>
<td>41.37</td>
</tr>
<tr>
<td>NTN4</td>
<td>mRNA</td>
<td>Down</td>
<td>6</td>
<td>34.588</td>
</tr>
<tr>
<td>SIK1</td>
<td>mRNA</td>
<td>Down</td>
<td>5</td>
<td>31.116</td>
</tr>
<tr>
<td>HECTD2</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>30.801</td>
</tr>
<tr>
<td>EPHA7</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>29.436</td>
</tr>
<tr>
<td>FRMD6</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>28.987</td>
</tr>
<tr>
<td>RND3</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>28.796</td>
</tr>
<tr>
<td>CFL2</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>27.316</td>
</tr>
<tr>
<td>RNF38</td>
<td>mRNA</td>
<td>Down</td>
<td>4</td>
<td>26.19</td>
</tr>
<tr>
<td>RBBP7</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.813</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.73</td>
</tr>
<tr>
<td>TMEM100</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.676</td>
</tr>
<tr>
<td>ANO6</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.587</td>
</tr>
<tr>
<td>DYPYSL2</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.573</td>
</tr>
<tr>
<td>STAR4D</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.292</td>
</tr>
<tr>
<td>RBMS1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.215</td>
</tr>
<tr>
<td>HLF</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.042</td>
</tr>
<tr>
<td>MMP2</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.019</td>
</tr>
<tr>
<td>GJA1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>25.018</td>
</tr>
<tr>
<td>CALD1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.922</td>
</tr>
<tr>
<td>TGFBI11</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.912</td>
</tr>
<tr>
<td>SYTL4</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.902</td>
</tr>
<tr>
<td>TSHZ3</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.844</td>
</tr>
<tr>
<td>FOXF1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.817</td>
</tr>
<tr>
<td>CLIP4</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.696</td>
</tr>
<tr>
<td>CYBRD1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.69</td>
</tr>
<tr>
<td>CNN1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.508</td>
</tr>
<tr>
<td>AHNAK</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.488</td>
</tr>
<tr>
<td>TIMP2</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.318</td>
</tr>
<tr>
<td>NBL1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>24.293</td>
</tr>
<tr>
<td>APCDD1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>23.427</td>
</tr>
<tr>
<td>SMOC1</td>
<td>mRNA</td>
<td>Down</td>
<td>3</td>
<td>23.331</td>
</tr>
<tr>
<td>PDGFRα</td>
<td>mRNA</td>
<td>Down</td>
<td>2</td>
<td>18.842</td>
</tr>
</tbody>
</table>

EPC, Edge Percolated Component; miRNA, microRNA.
Figure 3 GO and KEGG enrichment analyses of co-DEGs in prostate cancer. (A) Co-DEGs GO analysis; (B) Co-DEGs KEGG analysis; (C) enrichment relationship between co-DEGs and KEGG pathway. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genome; co-DEGs, common differentially expressed genes.

Figure 4 Differential expression and survival analyses of the hub targeted genes in prostate cancer obtained from TCGA and GTEx databases. *, P<0.05. TCGA, The Cancer Genome Atlas; GTEx, Genotype-Tissue Expression; PRAD, prostate adenocarcinoma; num, number; T, tumor; N, normal.
Figure 5 Independent predictive significance of a risk score in the TCGA and GSE70769 dataset, respectively. (A) Univariate and multivariate Cox regression analyses of clinic-pathological features and risk score in both sets. (B) Nomogram based on risk score and Gleason to predict BCRFS in the TCGA dataset at 1-, 3-, and 5-year. (C) Nomogram calibration plots predicting 1-, 3-, and 5-year BCRFS to evaluate the accuracy of candidate nomogram in both sets. (D) Time-dependent BCRFS ROC plots for TCGA and GSE70769 cohort. (E) Kaplan-Meier survival curves for biochemical relapse-free based on risk score in both sets. TCGA, The Cancer Genome Atlas; BCRFS, biochemical relapse-free survival; ROC, receiver operating characteristic; pN, pathological N staging; pT, pathological T staging; cM, clinical M staging; cT, clinical T staging; AUC, area under the curve.
prostate tissue. These genes were strongly associated with either overall survival (OS) or DFS, with good value for clinical diagnosis and prediction.

In this study, Gleason is an independent prognostic factor for biochemical recurrence in prostate cancer patients, which is consistent with previous studies (5,32). This novel nomogram based on two independent risk factors may be accurate for predicting 1-, 3- and 5-year BCRFS in patients with PCa, and its AUC is superior to that under the risk score alone or combined Gleason model.

This study has several limitations. Firstly, this study was a retrospective study using public databases. Secondly, the conclusion from this study needs to be validated with a large number of clinical samples.

In conclusion, our study suggests that this novel nomogram may be used as a biomarker for predicting biochemical recurrence and prognosis in patients with prostate cancer, with better sensitivity and specificity than existing clinical indicator models. Multi-center, prospective studies with large-sample sizes can be carried out in the future to verify our conclusion and clinical predictive value of this model.

Acknowledgments

Funding: This study was supported by the National Key Research and Development Program of China (No. 2020YFC2005404); the Scientific research and Cultivation program of Haidian District health development (No. HP2022-19-506005); the Youth Fund of Beijing Shijitan Hospital (No. 2020-q06).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-653/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-653/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted based on the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


Cite this article as: Su Q, Dai B, Zhang S. Construction of miRNA-mRNA network and a nomogram model of prognostic analysis for prostate cancer. Transl Cancer Res 2022. doi: 10.21037/tcr-22-653

Su et al. miRNA-mRNA network and a nomogram model in PCa