Abstract. Prostate cancer (PCa) is the second most common cancer amongst males worldwide. In the current study, microarray datasets GSE3325 and GSE6919 from the Gene Expression Omnibus database were screened to identify candidate genes that are associated with the progression of PCa. A total of 273 differentially expressed genes (DEGs) were identified, which included 173 downregulated genes and 100 upregulated genes, and a protein-protein interaction network was constructed using Search Tool for the Retired of Interacting Genes. The enriched functions and pathways of the identified DEGs included cell adhesion, the negative regulation of cell proliferation, protein binding and focal adhesion. A total of 8 hub genes were identified, of which PDZ binding kinase, Krüppel-like factor 4, collagen type XII α-1 chain, RAP1A and RAP39B were indicated to be associated with the progression and recurrence of PCa. In conclusion, the DEGs and hub genes identified in the present study may aid in determining the molecular mechanisms associated with PCa carcinogenesis and progression.

Introduction

Prostate cancer (PCa) is the second most common type of cancer among men globally, and constitutes ~15% of all cancer diagnoses worldwide (1). Digital rectal examination, measurement of the serum level of prostate specific antigen (PSA) and biopsy from a prostate transrectal ultrasonography are the most common diagnostic tools for PCa (2). Additionally, with advances in genetic analysis, alterations have been identified in a number of gene regions in patients with PCa, including prostate antigen 3, androgen-dependent transmembrane serine 2 and S-transferase P1 (3-5). However, genetic analysis exhibits a low specificity and can increase the number of unnecessary biopsies performed without reducing patient mortality (6). Previous studies have associated the tribbles pseudokinase 1 gene with the development of a number of tumors, including colorectal leukemia and hepatocellular cancers (7-9). It has been shown that transmembrane protease, serine 2:ETS-related gene (TMPRSS2:ERG) fusion is associated with diagnosing PCa in urine samples and DNA-based molecular templates (10). However, due to the lack of effective diagnostic methods during the early stages of the disease, the mortality rate of PCa remains high (10). Therefore, it is crucial to understand the molecular mechanisms associated with PCa carcinogenesis, proliferation and recurrence.

Microarray technology and bioinformatics analysis led to the identification of 273 differentially expressed genes (DEGs) and functional pathways in the carcinogenesis and progression of PCa. In the current study, two mRNA microarray datasets from Gene Expression Omnibus (GEO) were analyzed to identify DEGs between PCa tissues and non-cancerous tissues. Subsequently, the molecular mechanisms of PCa carcinogenesis and progression were investigated using Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and protein-protein interaction (PPI) network analyses. In conclusion, a total of 273 DEGs and 8 hub genes were identified in the current study, and these genes may be candidate biomarkers for PCa.

Materials and methods

Database. GEO (http://www.ncbi.nlm.nih.gov/geo) (11) is a public functional genomics database. GSE3325 (12) and GSE6919 (13) were downloaded from GEO (Affymetrix Human Genome U133 Plus 2.0 Array). The GSE3325 dataset contained 12 PCa tissue samples and 12 non-cancerous samples. GSE6919 contained 8 PCa samples and 8 non-cancerous samples.

Identification of DEGs. The Affy package (version 1.52.0) (14) was used to preprocess the raw expression data in the R statistical software (R x64 3.5.3; https://cran.r-project.org). DEGs were subsequently identified between PCa and normal samples using the limma (version 3.34.7) package of the R statistical software (https://bioconductor.org/packages/release/bioc/html/limma.html). DEGs with log2FC >1 and P<0.01 were selected in the microarray data.
Enrichment analysis of DEGs. The Database for Annotation, Visualization and Integrated Discovery (DAVID; version 6.7; http://david.ncifcrf.gov) (15), which provides functional annotation information of genes and proteins, was used to perform DEG analysis of KEGG pathway enrichment (16) and GO annotation (17). P<0.05 was set as the threshold value.

Module analysis and construction of the PPI network. The Search Tool for the Retrieval of Interacting Genes (version 10.0; http://string-db.org) (18), which offers comprehensive information on PPIs, was used to create the PPI network. The molecular interaction networks were visualized using Cytoscape (version 3.4.0; https://cytoscape.org/) (19). The Molecular Complex Detection (MCODE) (version 1.4.2) of Cytoscape was used to identify densely connected regions (20). The PPI networks were visualized using Cytoscape software. The hub genes and the most significant module in the PPI networks were identified using MCODE.

Hub genes selection and analysis. The hub genes were selected and their co-expression analysis were analyzed using cBioPortal (http://www.cbioportal.org) (21). Hierarchical clustering of hub genes was constructed using the University of California, Santa Cruz Cancer Browser (http://genome-cancer.ucsc.edu) (22). The overall survival and disease-free survival analyses of hub genes (the cutoff was the median expression value) were performed using Kaplan-Meier analysis in cBioPortal. The hazard ratio (HR) with 95% confidence intervals and log-rank P-value was also computed. The expression of PDZ binding kinase (PBK) and Krüppel-like factor 4 (KLF4) in cancer tissues were analyzed and presented using the online database Serial Analysis of Gene Expression (SAGE; http://www.ncbi.nlm.nih.gov/SAGE) (23). The relationship between expression patterns and tumor-node-metastasis (TNM) stage, Gleason grade and recurrence status were analyzed using the online database Oncomine (http://www.oncomine.com) (24).

Results

Identification of DEGs in PCa. After standardization of the microarray results, a total of 1,024 DEGs in GSE6919 and 2,371 DEGs in GSE3325 were identified. The overlap between the 2 datasets contained 273 genes, as presented in Fig. 1A, consisting of 173 downregulated genes and 100 upregulated genes in PCa tissues.

PPI network and module analysis. The PPI network of DEGs (Fig. 1B) and the most significant module were identified using Cytoscape (Fig. 1C). The functional analyses of DEGs demonstrated that genes in this module were mainly enriched in nucleotide binding, small molecule binding, focal adhesion and the regulation of the actin cytoskeleton (Table I).

Functional enrichment analyses of DEGs. Functional and pathway enrichment analyses of DEGs were performed using DAVID. GO analysis revealed that the biological processes of DEGs were significantly enriched in cell adhesion, negative regulation of cell proliferation, cell division and extracellular matrix organization (Table II). Molecular functions of DEGs were enriched in protein binding, GTP binding, mannose-binding, chromatin binding and chromatin binding (Table II). Cell components enriched with DEGs included the nucleus, cytoplasm, perinuclear region of the cytoplasm and focal adhesion (Table II). KEGG pathway analysis revealed that DEGs were mainly enriched in focal adhesion, regulation of the actin cytoskeleton, tight junction, coagulation cascades and gap junctions.

Analyses of the 8 hub genes. In the present study, a total of 8 hub genes were identified and these hub genes were presented in Table III. The criteria for selection were as follows: MCODE scores >5, degree cut-off=2, node score cut-off=0.2. Max depth=100 and k-score=2. Among the 8 genes, PBK, RAP1A, GNAS and RAB39B were upregulated, while COPZ2, KLF4, BACE1 and COL12A1 were downregulated. A network of the hub genes and their co-expression genes were analyzed using the cBioPortal online platform (Fig. 2A). Hierarchical clustering demonstrated that the hub genes could differentiate PCa samples from noncancerous samples (Fig. 2B). Subsequently, the overall survival analysis of the hub genes was performed using a Kaplan-Meier curve analysis. Patients with PCa and PBK, RAP1A, GNAS, coatomer protein complex subunit ζ 2 (COPZ2), β-secretase 1 (BACE1) and collagen type XII α-1 Chain (COL12A1) upregulation demonstrated decreased overall survival (Fig. 3A). Patients with PCa and PBK, RAP1A, GNAS, COPZ2, BACE1 and COL12A1 upregulation exhibited decreased disease-free survival (Fig. 3B). Additionally, RAB39B and KLF4 upregulation was associated with increased overall survival and disease-free survival. Based on the above survival analysis, PBK and KLF4 were identified to serve important roles in the carcinogenesis or progression of PCa. Oncomine analysis of cancer and normal tissue revealed that PBK and KLF4 were significantly over-expressed in PCa in the different datasets (Fig. 4A and B). In the Taylor prostate of Oncomine dataset, the increased mRNA levels of PBK were associated with TNM stage, Gleason grade and recurrence status (Fig. 5A). In the Tatulippe prostate of Oncomine dataset, decreased KLF4 mRNA levels were associated with TNM stage, Gleason grade and recurrence status (Fig. 5A and B). PBK gene expression in metastatic tissue was higher compared with primary tumor and solid tissue normal via Oncomine (Fig. 6).

Discussion

Previous studies have demonstrated that the TMPRSS2:ERG fusion is significantly associated with the diagnosis of PCa (25-27). Promoter hypermethylation and downregulated expression of glutathione peroxidase 3 have been observed in a variety of cancer types, including thyroid cancer, hepatocellular carcinoma and PCa (10,26,27). Yu et al (28) identified an association between Piwi-like protein 2 (PIWIL2) gene expression and metastatic PCa. Potential markers for use in the diagnosis and treatment of PCa, which exhibit high efficiency, are urgently required. To increase understanding of the molecular mechanisms of candidate genes, GO, KEGG and PPI analyses were performed. In the current study, the epigenetic and genetic mechanisms in PCa were assessed using microarray technology.
A total of two mRNA microarray datasets were selected. A total of 273 DEGs were identified, including 173 downregulated genes and 100 upregulated genes. The interactions of DEGs were investigated using GO and KEGG analyses. DEGs were found to be enriched in focal adhesion, regulation of the actin cytoskeleton, tight junctions, coagulation cascades and gap junctions. However, other studies (29,30) have demonstrated that DEGs were enriched in a number of functional terms, including cellular response to bone morphogenetic protein (BMP) stimulus, response to BMP, extracellular region and pathways that are associated with transforming growth factor-β signaling. GO enrichment analysis revealed that changes in the most significant modules were enriched in nucleotide binding, nucleoside phosphate binding, small molecule binding and cytoskeletal protein binding, while changes in KEGG were mainly enriched in focal adhesion and regulation of the actin cytoskeleton.

A total of 8 DEGs were selected as hub genes. The criteria for selection were as follows: MCODE scores >5, degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score=2. Among the 8 genes, PBK, RAP1A, GNAS and RAB39B were upregulated, while COPZ2, KLF4, BACE1 and COL12A1 were downregulated. PBK and KLF4 were identified to be important genes in the present study. PBK is highly homologous to mitogen-activated protein kinase (31,32). By virtue of target utilization, PBK has been revealed to influence growth and differentiation (33-36). PBK is expressed in the outer cell layer of seminiferous tubules in primary spermatocytes (37), and is often increased in a number of human cancer types from different tissue types.

<table>
<thead>
<tr>
<th>Pathway ID</th>
<th>Pathway description</th>
<th>Count in gene set</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0000166</td>
<td>Nucleotide binding</td>
<td>6</td>
<td>0.003</td>
</tr>
<tr>
<td>GO:1901265</td>
<td>Nucleoside phosphate binding</td>
<td>6</td>
<td>0.003</td>
</tr>
<tr>
<td>GO:0036094</td>
<td>Small molecule binding</td>
<td>6</td>
<td>0.004</td>
</tr>
<tr>
<td>GO:0008092</td>
<td>Cytoskeletal protein binding</td>
<td>5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>hsa04510</td>
<td>Focal adhesion</td>
<td>8</td>
<td>0.007</td>
</tr>
<tr>
<td>hsa04810</td>
<td>Regulation of actin cytoskeleton</td>
<td>8</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Table I. GO and KEGG pathway enrichment analysis of DEGs in the most significant module.

| GO:0007155     | Cell adhesion                     | 12                | 0.049  |
| GO:0008285     | Negative regulation of cell proliferation | 11        | 0.044  |
| GO:0051301     | Cell division                     | 10                | 0.050  |
| GO:0030308     | Negative regulation of cell growth | 7                 | 0.006  |
| GO:0030198     | Extracellular matrix organization  | 7                 | 0.051  |
| GO:0005634     | Nucleus                           | 99                | <0.001 |
| GO:0005737     | Cytoplasm                         | 93                | <0.001 |
| GO:0048471     | Perinuclear region of cytoplasm   | 17                | 0.010  |
| GO:0005925     | Focal adhesion                    | 16                | <0.001 |
| GO:0009986     | Cell surface                      | 14                | 0.033  |
| GO:0005515     | Protein binding                   | 147               | <0.001 |
| GO:0005525     | GTP binding                       | 13                | 0.007  |
| GO:0003682     | Chromatin binding                 | 12                | 0.019  |
| GO:0019901     | Protein kinase binding            | 11                | 0.035  |
| hsa04510       | Focal adhesion                    | 11                | 0.007  |
| hsa04810       | Regulation of actin cytoskeleton  | 9                 | 0.008  |
| hsa04530       | Tight junction                    | 6                 | 0.003  |
| hsa04540       | Gap junction                      | 6                 | 0.005  |

Table II. KEGG and GO enrichment analyses of DEGs.

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEG, differentially expressed genes; FDR, false discovery rate.
However, the function of PBK has not yet been fully determined. In a previous study, the immunohistochemical expression of PBK/T-LAK cell-originated protein kinase (TOPK) was revealed to be significantly associated with human bladder cancer, and was identified as a novel diagnostic biomarker for this disease (40). In the present

Table III. Functional roles of 8 hub genes with degree ≥10.

<table>
<thead>
<tr>
<th>No.</th>
<th>Gene symbol</th>
<th>Full name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PBK</td>
<td>PDZ binding kinase</td>
<td>Active lymphoid cells and support testicular functions; over expression of this gene has been implicated in tumorigenesis</td>
</tr>
<tr>
<td>2</td>
<td>RAP1A</td>
<td>Member of RAS oncogene family</td>
<td>Affect cell proliferation and adhesion, and may play a role in tumor malignancy</td>
</tr>
<tr>
<td>3</td>
<td>GNAS</td>
<td>Member of RAS oncogene family</td>
<td>The encoded protein regulates signaling pathways that affect cell proliferation and adhesion</td>
</tr>
<tr>
<td>4</td>
<td>COPZ2</td>
<td>Coatomer protein complex subunit ζ 2</td>
<td>This gene encodes a member of the adaptor complexes small subunit family</td>
</tr>
<tr>
<td>5</td>
<td>KLF4</td>
<td>Krüppel-like factor 4</td>
<td>Control the G1-to-S transition of the cell cycle</td>
</tr>
<tr>
<td>6</td>
<td>BACE1</td>
<td>β-secretase 1</td>
<td>This gene encodes a member of the peptidase A1 family of aspartic proteases</td>
</tr>
<tr>
<td>7</td>
<td>COL12A1</td>
<td>Collagen type XII α 1 chain</td>
<td>Modify the interactions between collagen I fibrils and the surrounding matrix</td>
</tr>
<tr>
<td>8</td>
<td>RAB39B</td>
<td>Member RAS oncogene family</td>
<td>Encodes a member of the Rab family of proteins</td>
</tr>
</tbody>
</table>

Figure 1. Venn diagram, PPI network and the most significant module of DEGs. (A) DEGs were selected with a fold change >2 and P-value <0.01. The 2 datasets showed an overlap of 273 genes. (B) PPI network of DEGs constructed using Cytoscape. (C) Most significant module with 20 nodes and 100 edges. Upregulated genes are marked in light red; downregulated genes are marked in light blue. DEGs, differentially expressed genes; PPI, protein-protein interaction.
Figure 2. Interaction network and biological process analysis of the hub genes. (A) Hub genes and co-expression genes were analyzed using cBioPortal. Nodes with bold black outline represent hub genes. Nodes with thin black outline represent the co-expression genes. (B) Hierarchical clustering of hub genes. The samples under the red bar are non-cancerous samples and the samples under the blue bar are PCA samples.

Figure 3. Correlation of expression of individual hubs genes survival. (A) Overall survival and (B) disease-free survival analyses of hub genes were performed using cBioPortal online platform. $P<0.05$ was considered statistically significant.
Figure 4. Oncomine analysis of cancer vs. normal tissue of PBK and KLF4. Heat maps of (A) PBK and (B) KLF4 gene expression in clinical PCA carcinoma samples vs. normal tissues. PBK, PDZ binding kinase; KLF4, Krüppel-like factor 4.

Figure 5. Relationship between the PBK, KLF4 and clinicopathological features. Association between the expression of (A) PBK and (B) KLF4 and TNM stage, Gleason grade and recurrence status. PBK, PDZ binding kinase; KLF4, Krüppel-like factor 4; TNM, tumor-node-metastasis.
PBK, a member of the protein kinase family, plays a crucial role in regulating cellular proliferation, differentiation, and apoptosis. Its overexpression is linked to tumor cell proliferation, and it has been shown to promote tumor cell survival and disease-free survival. PBK expression is significantly increased in PCa, with higher mRNA levels associated with TNM stage, Gleason grade, and recurrence status. It also serves as an independent predictor of PCa recurrence and is negatively associated with overall and disease-free survival.

KLF4, a member of the Krüppel-like zinc finger transcription factor family, is also associated with survival of no biochemical recurrence. KLF4 protein is downregulated in PCa tissue with metastases, and its expression is significantly decreased in the Gleason high-scoring group compared with the control group. This indicates that KLF4 can serve as an oncogene or tumor suppressor gene in PCa. KLF4 targeting genes are also biomarker transcription factors in the endothelial-mesenchymal transition (EMT) process. Additionally, a previous study indicated that the expression of E-cadherin and α-catenin in the KLF4 overexpression treatment group was significantly higher compared with the control group, while the mesenchymal cell marker vimentin and the expression of vascular endothelial growth factor were significantly lower compared with the control group. It has been shown that KLF4 protein is negatively associated with clinical stages in patients with meningioma, and it promotes or inhibits the EMT process by acting on transcription factors. The transcription factor KLF4 in PCa cells promotes the migration and invasion of EMT and tumor cells in vitro. These results are consistent with the results of the current study, which indicated that lower mRNA levels of PBK were associated with TNM stage, Gleason grade, and recurrence status. KLF4 was also indicated to be downregulated in PCa tissue with metastases. Furthermore, the stable knockdown of KLF4 expression in PCa cells has been identified to upregulate the expression of epithelial-related gene E-cadherin and downregulate the expression of a variety of mesenchymal-associated genes in vitro, and has been revealed to serve a role in the inhibition of tumor cell migration and invasion. Katz et al demonstrated that the expression of KLF4 in tumor tissues was significantly decreased in patients with PCa in the USA, and that the upregulation of KLF4 inhibited tumor migration and invasion. Ghaleb et al identified a positive feedback loop control between KLF4 and the androgen receptor, and revealed that the inhibition of KLF4 expression in prostate adenocarcinoma cells can inhibit the occurrence of EMT in vitro and serve a role in inhibiting tumor cell migration and invasion. It has also been indicated that KLF4 can serve the role as an oncogene or tumor suppressor gene in a number of cellular environments.

In conclusion, a total of 273 DEGs and 8 hub genes were identified as potential novel diagnostic biomarkers for PCa. The current study identified 2 genes associated with PCa progression, including PBK and KLF4. However, the current study is performed based on bioinformatics methods and no experiments were performed to confirm these conclusions. Therefore, further experimental study is required to support the results gained from the current analysis.
et al. The cBio: TRIB1 promotes colorectal cancer cell

References

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

Funding

The present study was supported by Special Projects for Key

Availability of data and materials

The datasets used and/or analyzed during the present study are

Authors' contributions

SL, WBX and JH conceived and designed the study, and

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

1. Ferlay J, Soerjomataram I, Dikshit R, Eser M, Mathers C, 
   Rebelo M, Parkin DM, Forman D and Bray F: Cancer incidence and
   mortality worldwide: sources, methods and major patterns in

   Bergh RC, Bangma CH and Schröder FH: A risk-based strategy
   improves prostate-specific antigen-driven detection of prostate

3. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smitt FP,
   Karthaus HF, Schalken JA, Debruyne FM, Ro JN and Isaacs WB:
   DD3: A new prostate-specific gene, highly overexpressed in

4. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R,

5. Harden SV, Sanderson H, Goodman SN, Partin AA, Walsh PC,


