Identification of hub genes and their clinical value for predicting the development of prostate cancer from benign prostate hyperplasia by bioinformatic analysis

Xi Chen  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University  
https://orcid.org/0000-0002-5674-9446

Junjie Ma  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Chengdang Xu  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Licheng Wang  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Yicong Yao  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Xinan Wang  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Tong Zi  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Cuidong Bian  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Denglong Wu  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University

Gang Wu  
Department of Urology, Tongji Hospital, School of Medicine, Tongji University  
https://orcid.org/0000-0001-8669-4944

Primary research

Keywords: Benign prostate hyperplasia, Prostate cancer, Differentially expressed genes, Clinical value, Bioinformatic analysis

Posted Date: September 9th, 2021
Abstract

Background

Prostate cancer (PCa) and benign prostate hyperplasia (BPH) are commonly encountered diseases in elderly males. The two diseases have some commonalities: both are growth depend on hormone and respond to antiandrogen therapy. Some studies have shown that genetic factors are responsible for the occurrences of both diseases. There may be a correlation between BPH and PCa.

Methods

The GEO database can help determine the differentially expressed genes (DEGs) between BPH and PCa. Gene Ontology (GO) term enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were utilized to find pathways in which the DEGs were enriched. The STRING database can provide a protein–protein interaction (PPI) network, and Cytoscape software can find hub genes in PPI network. GEPIA can be used to analyze expression and survival data for hub genes. R software was used to progress regression analysis, decision curve analysis and built nomograph. UALCAN and The Human Protein Atlas was utilized to test the results. Finally, we made clinical and cell experiments to verify the results.

Results

Sixty DEGs, consisting of 15 up-regulated and 45 down-regulated genes, were found based on the GEO database. Using Cytoscape, we found 7 hub gene in the PPI network. The hub gene expression was tested on TCGA database. Except CXCR4, all hub genes expressed differently between tumor and normal samples. Meanwhile, all hub genes exclude CXCR4 has diagnostic value in predicting PCa and their mutations are risk factors leading to PCa. The expression of CSRP1, MYL9 and SNAI2 changed in different tumor stage. CSRP1 and MYH11 could affect the disease-free survival (DFS). The same results reflected in different database. In addition, we also chose three hub gene, MYC, MYL9, and SNAI2, to validate their functions in clinical specimens and cells.

Conclusion

These identified hub genes can help us to understand the process and mechanism by which BPH develops into PCa and provide achievable targets for predicting which BPH patients may later develop PCa.

Background

Benign prostate hyperplasia (BPH) is common among elderly men over 70. Prostate cancer (PCa) is still have high incidence rates of all cancers and is a major reason to death in elderly men[1]. Although BPH and PCa are different diseases, in that BPH is a benign disease that arises in the transitional zone, and PCa is a malignant tumor that arises mainly in the peripheral zone, they are related. The relationship
between BPH and PCa was first noted in the 1950s in studies of prostate glands. During the past 60 years, some studies have shown that BPH and PCa have some definite associations. Sommers first studied BPH and PCa on cadavers. In their work, BPH was found in 80% and 45% of cadavers with or without PCa, respectively[2]. Later, another study covered alike results[3]. In 2002, Hammarsten and Hogstedt reported that BPH which has a faster growing speed may be hazards for developing clinical PCa[4]. Another study proved that the volume of the prostate may be one of the reasons for the aggressiveness of PCa, and PCa located in small glands is more aggressive than that located in larger glands[5]. This means that BPH may affect the degree of malignancy of PCa. Orsted et al. followed 3,009,258 Danish men from 1980 to 2006. During 27 years of follow-up, they found that clinical BPH was linked with a raised risk of PCa and a higher risk of death by BPH[6].

In 2001, Luo et al. researched the genetic relationship between BPH and PCa, and found that 3,215 genes were expressed differently between BPH and PCa samples[7]. Some studies have reported that gene expression could be a causal factor in the development of PCa from BPH and may even affect the degree of malignancy of PCa[8, 9]. Microarray technology and bioinformatic analysis have been extensively wielded to analyze differentially expressed genes (DEGs) and can be used to find functional pathways that will help us to better understand PCa[10]. Through gene expression profiling, some investigations have found many DEGs that play critical roles in the process of the development of PCa from BPH[8, 9]. However, dependence on a single microarray analysis may lead to inaccurate results. Therefore, examination of different microarray analyses can provide more reliable results.

The Gene Expression Omnibus (GEO) database can provide many microarray datasets. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and protein–protein interaction (PPI) network analyses have been performed to help us understand the potential mechanisms for the occurrence and progression of diseases. Thus, we used bioinformatic analysis to find key genes that may be important for development of PCa from BPH. Then, based on genetic and clinical data from The Cancer Genome Atlas (TCGA) database, the diagnostic model was built to predict the hub genes of diagnostic value for PCa. Tumor staging and survival time were also analyzed. Logistic regression was used for predicting the mutation of hub genes leading to PCa. The UALCAN and The Human Protein Atlas databases could test the expression of hub gene in normal specimens and PCa tissues. Finally, we validated the function of three DEGs, MYC, MYL9, and SNAI2, in PCa based on clinical specimens and C4-2 PCa cells.

Materials And Methods

Microarray data

The GEO database (http://www.ncbi.nlm.nih.gov/geo/) at the National Center for Biotechnology Information (NCBI) is a communal database that provides a genomics data repository of gene expression, chip, and microarray data[11]. The criteria for GSE data included in the study as follow: 1. The GSE samples have complete gene expression data from high-throughput sequencing and can be
downloaded from GEO database. 2. The GSE samples data included both BPH samples and PCa samples. 3. There is a clear definition which are BPH samples and which are PCa specimens. Then we found three datasets GSE5377, GSE104749 and GSE30994 met our criteria. Then we downloaded the three datasets from GEO database. GSE5377 (based on the GPL 201 Affymetrix Human HG-Focus Target Array) included 3 BPH samples and 17 PCa samples. GSE104749 (based on the GPL 570 Affymetrix Human Genome U133 Plus 2.0 Array) included 4 BPH samples and 4 PCa samples. GSE30994 (based on the GPL4133 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F [Feature Number version]) included 3 BPH samples and 3 PCa samples. Overall, 10 BPH and 24 PCa samples were enrolled in our study.

**Data handling and DEGs searching**

The primary data were got and normalized by R language. According to comments of the documents, the expression matrix including probe ID was substituted by the corresponding gene ID, and if there were multiple probes that corresponded to the same gene, the average value was calculated using the R software for further study[12]. Then all genes of each data set were searched using the limma R package, and genes with an adjusted P-value < 0.05 and |log2fold change (FC)|>1 were considered DEGs. Then, we used the online web tool, Venn diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/) to find the integrated DEGs. In addition, the up-regulated and down-regulated genes were downloaded for further study.

**GO and KEGG pathway analysis of DEGs**

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 (https://david.ncifcrf.gov/) was used to perform GO functional and KEGG pathway analyses of the integrated DEGs[13]. The GO functional analysis of integrated DEGs involved three parts: biological processes (BP), cell components (CC), and molecular functions (MF). P < 0.05 was considered statistically statistical differences [14].

**PPI network and module analysis**

A PPI network of the integrated DEGs was structured by online tool the Search Tool for the Retrieval of Interacting Genes (STRING) database with the default medium confidence (0.4) (http://www.string-db.org/)[15]. It helped us to find the key genes and critical gene modules participated in the promotion of BPH to PCa. Cytoscape software was used for reconstructing the PPI network, and module and GO analyses were carried out by two plug-ins in Cytoscape, Molecular Complex Detection (MCODE) and Biological Network Gene Ontology tool (BiNGO), to clarify the biological significance of gene modules from BPH to PCa. P < 0.05 indicated a significant difference, and these genes were designated as hub genes[16].

**Construction of risk prediction model and survival analysis**

Hub gene expression between normal prostate specimens and PCa tissues was compared using gene expression profiling interactive analysis (GEPIA; http://www.gepia.cancer-pku.cn/) dependent on TCGA
database[17]. Logistic regression was performed to screen the hazard ratios of hub genes changes leading to PCa. A nomogram was built to predict the risk value of the hub genes. A forest map was utilized to show the hazard ratios more intuitively. Moreover, the prognostic value of gene was enucleated by GEPIA. Then, overall survival (OS), disease-free survival (DFS) was analyzed too.

**Construction of the diagnostic model and decision curve analysis**

To further analyze the hub genes’ diagnostic value for PCa, we collected the gene expression of hub genes and clinical data from TCGA databases (https://portal.gdc.com)[18]. After all data were collected, we used GraphPad Prism 7 (GraphPad Software, Inc., San Diego, CA) to draw the receiver operating characteristic (ROC) curve to reflect the accuracy of the model. Then decision curve analysis (DCA) was carried out by R software.

**Expression of hub genes at different tumor stages**

Tumor-Node-Metastasis (TNM) classification of malignant tumors is commonly used to assess the tumor severity. American Joint Commission on Cancer (AJCC) has another classification on tumor[19]. Study the relationship between altered gene expression and tumor can help understanding the effect of genetic changes on tumor progression. Clinical data such as TNM classification and ACJJ tumor classification was got from TCGA database. Then, we analyzed the hub genes’ expression correlation with these clinical data.

**Validation of hub genes based on different databases**

Because the normal tissue samples from GEPIA including sequencing from both TCGA and GTEx database. So, we further analyzed the hub genes’ expression in normal prostate samples and PCa specimens on the UALCAN database (http://ualcan.path.uab.edu/) and The Human Protein Atlas (https://www.proteinatlas.org/)[20, 21].

**Clinical specimen validation**

Quantitative real-time PCR (RT–qPCR), Western blotting (WB), and immunohistochemistry (IHC) were used for analyzing clinical specimens. The total RNA was extracted from tumor and para-cancerous tissues of 6 PCa patients who were previously diagnosed with BPH utilizing TRIzol Reagent (Sigma–Aldrich, St. Louis, MO, USA). cDNA was transcribed using the reverse transcription kit (Advantage® RT-for-PCR Kit, Takara Bio Inc., Kusatsu, Japan). Finally, we measured the volume of cDNA using real-time PCR reagents and a kit (TB Green® Premix Ex Taq™ II, Takara Bio Inc.) according to the manufacturer’s descriptions. The following primers of c-MYC, MYL9, SNAI2 and GAPDH were shown in Table 1. The $2^{-\Delta\Delta Ct}$ method was used to quantify mRNA expression levels.
Table 1
Primers used for the qRT-PCR

| Gene Name | Primer sequence                     |
|-----------|-------------------------------------|
| c-MYC     | Forward: 5’-AGCGACTCTGAGGAGGAACAA-3’ |
|           | Reverse: 5’-TGGGCTGTGAGGAGGTTTG-3’  |
| MYL9      | Forward: 5’-AACATGTCCAGCAAACGTGC-3’ |
|           | Reverse: 5’-GCGAAGACATTGGAGGTGG-3’  |
| SNAI2     | Forward: 5’-GGACTAGTAGTAGGCCGCTCCTTCTCTGGTC-3’ |
|           | Reverse: 5’-CGGAATTCTCAGTGTGCTACACAGCAGCCAGATTC-3’ |
| GAPDH     | Forward: 5’-GGAGCGAGATCCCTCCAAAAT-3’ |
|           | Reverse: 5’-GGCTGTTGTCATACTTCTCATGG-3’ |

Protein was extracted with RIPA lysis buffer from 6 paired patients’ para-cancerous normal tissues and cancer tissues. Protein samples were treated with Dual Color Protein Loading Buffer (Thermo Fisher Scientific, Waltham, MA, USA). SDS–PAGE gels (10% and 15%) were used to separate proteins, followed by transfer to nitrocellulose membranes (NC) (Merck KGaA, Darmstadt, Germany). Protein-Free Rapid Blocking Buffer (Thermo Fisher Scientific) was utilized to block the membranes. Then the membranes were incubated overnight at 4°C with primary antibodies against c-MYC (1:1000), MYL9 (1:1000), and SNAI2 (1:1000) (Abcam UK, Cambridge, UK). The next day, 1xTBST was used to wash the membranes three times (10 min. each). Then, the membranes were incubated at room temperature for 1 h with a matched secondary antibody (HRP-labeled Goat Anti-Human IgG (H + L), Beyotime Biotechnology, Shanghai, China). Lastly, the membranes were exposed to X-ray film.

Then, the expression of MYC, MYL9, and SNAI2 in 2 clinical patients’ specimens was detected by IHC. Prostate sections were deparaffinized with xylene and then hydrated through a stepwise ethanol gradient. After antigen repair, endogenous peroxidase activity was blocked by incubation with 3% H₂O₂, and nonspecific binding was blocked using 5% goat serum. Sections were incubated with anti-MYC, anti-MYL9, and anti-SNAI2 antibodies (Abcam UK) at 4°C overnight and then with horseradish peroxidase-labeled secondary antibody. Slices were scanned, and images were processed using an Olympus microscope (Olympus Corp., Tokyo, Japan) and Olympus cellSens software (Olympus Corp.).

Validation in C4-2 prostate cancer cell lines

C4-2 PCa cells were transfected with a MYC breakdown plasmid, a MYL9 overexpression plasmid, and a SNAI2 overexpression plasmid (constructed by fenghbio company, Hunan, China). Then transfected C4-2 cells were used in invasion and proliferation experiments. Approximately 1*10⁵ C4-2 cells and 150 uL 2% fetal bovine serum (FBS) (gibco, Thermo Fisher Scientific, Waltham, MA, USA) + 1640 culture medium (sigma Darmstadt, Germany) was put in the upper chamber, and 10% FBS + 1640 culture medium was
placed in the lower cubicle. After 36h, cells were fixed with 4% paraformaldehyde fixative solution. The cells were stained with crystal violet and observed by an Olympus microscope (Olympus Corp.). ImageJ was utilized to count cell numbers. About 1000 C4-2 cells were placed in each well of a 96-well plate. Each set was repeated three times. The proliferation of cells in 0h, 24h, 48h and 72 h were detected by Cell Counting Kit-8 (CCK-8). The optical density (OD) at 450nm was measured by enzyme labeling.

**Statistical analysis**

The matrix data was handled with R version 4.0.2 (Institute for Statistics and Mathematics, Vienna, Austria; https://www.r-project.org) and the statistical analysis was finished by SPSS 25.0 (IBM Corp., Armonk, NY, USA). For descriptive statistics, mean ± standard deviation was used for continuous variables with normal distributions, whereas the median (range) was used for continuous variables with abnormal distributions. Categorical variables were described by counts and percentages. Hazard ratios (HRs), the 95% confidence interval (95% CI), and P values were used as statistical metrics. Two-tailed P < 0.05 was deemed as statistically significant.

**Results**

**Identification of DEGs**

Gene expression datasets GSE5377, GSE107479, and GSE30994 were acquired from the GEO database. All datasets were normalized by R language, and then the DEGs were screened using the limma R package (adjusted P < 0.05 and |log2fold change (FC)|>1). The GSE5377 dataset included 547 DEGs, with 167 up-regulated genes and 380 down-regulated genes. The GSE104749 dataset included 3790 DEGs consisting of 833 up-regulated genes and 2957 down-regulated genes. The GSE30994 dataset contained 3790 DEGs, including 1429 up-regulated genes and 1872 down-regulated genes. The up and down regulated DEGs was shown in Venn map. (Fig. 1)

**GO and KEGG pathway analysis of DEGs**

Try to find the pathways which DEGs enriched, we used DAVID to find the pathways. Up-regulated genes were enriched in genes involved in the regulation of the plasma membrane. The down-regulated genes were mainly enriched in genes involved in negative regulation of nitric oxide biosynthesis, regulation of extracellular exosome expression, and regulation of glutathione binding. In the KEGG pathway analysis, we also found that the down-regulated genes were mainly enriched in proteoglycans in cancer and focal adhesion. The KEGG pathways which DEGs enriched was shown in Fig. 2C (Fig. 2C). The GO analysis was shown in Fig. 2D. (Fig. 2D).

**PPI network and the most significant module analysis**

The PPI network of DEGs was constructed to help find the hub genes. The most significant module was obtained using Cytoscape software, and we found 7 hub genes. (Fig. 2A). Moreover, the contact of the 7
hub genes was also analyzed. (Fig. 2B)

**Expression of hub genes in TCGA database**

To further confirm that the hub genes could be important factors leading to PCa, we used the GEPIA database to compare the hub genes’ expression in normal samples and PCa specimens. As shown in Fig. 3, the hub genes were expressed differently in normal samples and PCa specimens, except for *MYC* and *CXCR4*. (Fig. 3) Then, we assessed the clinical data of all PCa patients in TCGA database. (Table 2)

| Project       | Category                        | Statistics Results |
|---------------|---------------------------------|--------------------|
| Gender        | Male                            | 490                |
| Age (Years)   |                                 | 61[41–78]          |
| Race          | White                           | 407(83.1%)         |
|               | Asian                           | 12(2.5%)           |
|               | Black                           | 58(11.8%)          |
|               | Not reported and Indian         | 13(2.6%)           |
| Survival      | Alive                           | 479(97.8%)         |
|               | Dead                            | 9(1.8%)            |
|               | Not reported                    | 2(0.4%)            |
| Tumor stage   | T2                              | 187(38.2%)         |
|               | T3                              | 286(58.4%)         |
|               | T4                              | 10(2%)             |
|               | Not reported                    | 7(1.4%)            |
| Node stage    | N0                              | 340(69.4%)         |
|               | N1                              | 78(15.9%)          |
|               | Not reported                    | 72(14.7%)          |
| Treatment     | Pharmaceutical therapy          | 244(49.8%)         |
|               | Radiation therapy               | 246(50.2%)         |

**Results of ROC curve and decision curve analyses**

The ROC curve was used to analyze the accuracy of each hub gene's ability to diagnose PCa. At the same time, it was also a process for verifying the above results. *CXCR4* had poor diagnostic efficacy, with an
Area Under Curve (AUC) of 0.5198 (95% CI, 0.4316–0.6079; P = 0.6419). The other hub genes had perfect diagnostic values (P < 0.05). MYC had an AUC of 0.7553 (95% CI, 0.6773–0.8332), CSRP1 had an AUC of 0.8764 (95% CI, 0.8294–0.9234), SNAI2 had an AUC of 0.7399 (95% CI, 0.7830–0.8968), MYL9 had an AUC of 0.8300 (95% CI, 0.7727–0.8873), ACTG2 had an AUC of 0.8499 (95% CI, 0.7956–0.9042), and MYH11 had an AUC of 0.8651 (95% CI, 0.8103–0.9198). (Fig. 4A and Table 3) This means that almost all the hub genes were meaningful for the diagnosis of PCa. In addition, decision curve analysis (DCA) is another model used in predicting diagnose value. We made DCA to evaluate all hub genes value in predicting PCa. We found that hub genes together have better diagnose value in predicting PCa. (Fig. 4B)

**Table 3**
The diagnostic value of hub genes in PCa

| Id    | P value | AUC and 95% CI       |
|-------|---------|----------------------|
| MYC   | 0.0027  | 0.7553(0.6773–0.8332)|
| CXCR4 | 0.6419  | 0.5198(0.4316–0.6079)|
| CSRP1 | 0.0009  | 0.8764(0.8294–0.9234)|
| SNAI2 | 0.0079  | 0.7399(0.7830–0.8968)|
| MYL9  | 0.0153  | 0.8300(0.7727–0.8873)|
| ACTG2 | 0.0178  | 0.8499(0.7956–0.9042)|
| MYH11 | 1.76E-07| 0.8651(0.8103–0.9198)|

**Hub gene expression at different tumor stages**

Using the clinical data downloaded from TCGA database, we then analyzed hub gene expression at different tumor stages. We analyzed the hub genes’ expression in different TNM classification and ACJJ classification of malignant tumors. We found that some hub genes expression will also change when PCa progression. For example, the expression of CXCR4 will increase significantly from T2 tumor stage to T3 tumor stage (P = 0.028). In addition, CSRP1 and MYL9 will decrease when tumor progressed from T2 to T4 (P = 0.032 and P = 0.047). The expression level of CXCR4 and SNAI2 will change when node metastasis happened (P = 0.022 and P = 0.012). In addition, MYC expressed differently in different ACJJ T1 and T2 stages (P = 0.034). (Fig. 5)

**Risk prediction model and survival analysis**

Logistic regression can be used to predict risk factors which can lead to disease. As a type of generalized linear model, logistic regression has been widely used in disease diagnosis. Here, we made logistic analysis to predict hub genes’ expression in resulting PCa. The nomogram was constructed to forecast the probability of hub gene mutation leading to PCa. (Fig. 6A) A calibration curve was made to verify the nomogram and compare the nomogram risk and actual risk. (Fig. 6B) Single factor and multi-factor regression showed the hub genes’ mutation risk in causing PCa. (Fig. 6C-D) In single factor logistic
regression forest map, we found that except CXCR4, \( P = 0.848 \) all other hub genes may be risk factors in the occurrence of PCa. However, in the multiple factor logistic regression forest map, we found that only SNAI2 \( (P = 0.04) \) and MYH11 \( (P = 0.024) \) may be risk factors in leading to PCa. The residuals plot and the normal P–P plot of standardized residuals of logistic regression were used to test the effect of the regression model. (Figure S1A-B) We further studied the relationship between patients’ survival and hub gene expression. Some hub genes like CSRP1 and MYH11 could affect the length of disease-free survival (DFS). (Fig. 7) However, no hub genes had an effect on overall survival time. (Figure S2)

**Validation using other databases**

To verify the above results, we used the UALCAN and The Human Protein Atlas databases to compare the hub genes’ expression in normal and tumor specimens. Like the results found above, the other 6 hub genes except CXCR4 were expressed differently in normal prostate tissue and PCa samples. (Fig. 8)

**Validation in clinical specimens**

Finally, we analyzed the hub genes’ expression in our clinical specimens. MYC and CXCR4 was up-regulated in tumor tissues than normal tissues However, CXCR4 expressed no difference in GEPIA and UALCAN. In addition, CXCR4 did not has diagnose value, we choose MYC, as the upregulation gene, for experiment validation. In addition, there are 5 hub genes down-regulated. All the down-regulated genes expressed differently in TCGA. In ROC curve analysis, we found that MYL9 and SNAI2 has the lowest AUC among 5 down-regulated hub genes. That means the two hub genes may have minimum authenticity among 5 down-regulated hub genes in predicting PCa. So, we finally choose MYC, MYL9 and SNAI2 as the target genes for further study. We included 6 patients who were diagnosed with PCa with previous diagnoses of BPH. We collected the cancer and para-cancerous tissues. We carried out RT–qPCR and Western blot experiments to compare the mRNA and protein expression of MYC, MYL9, and SNAI2. In addition, 2 paired samples from PCa patients were used in IHC experiment. We found that MYC was truly increased in both mRNA and protein level when PCa happened. (Fig. 9A-B) MYL9 expressed lower in tumor tissues than normal tissues at mRNA and protein level. (Fig. 9C-D). Meanwhile, SNAI2 was also down-regulated in PCa sample tissues than para-cancerous normal tissues. (Fig. 9E-F) At the same time, the IHC results reflect the same trend with the results from RT-qPCR and Western blot. (Fig. 9G-I)

**Hub genes influence on C4-2 prostate cancer cell lines**

To further verify whether the three hub genes we found can affect cell ability, we constructed shMYC plasmid, oeMYL9 and oeSNAI2 plasmid. We verify the mRNA and protein expression level after cell transfected plasmid. We successfully knock-down MYC in C4-2 cell. (Fig. 10A-B). The oeMYL9 and oeSNAI2 C4-2 cell was constructed successfully, too. (Fig. 10C-F) After experiment cell line constructed, we made transwell and CCK8 experiment to compare the invasion and proliferation ability between normal control cell line and experiment cell line. We found that after transfected by shMYC, oeMYL9 and oeSNAI2 plasmid, the C4-2 prostate cancer cell invasion ability will decrease. (Fig. 10G). The same results
reflect in CCK8 experiments, too. When C4-2 prostate cancer cell transfected with shMYC, oeMYL9 and oeSNAI2 plasmid, the cell proliferation will decrease. (Fig. 10H)

**Discussion**

PCa has a higher and higher incidence among all tumors in men, especially those over 70. BPH is another common encountered disease and affects nearly seventy percent of men older than 70[22]. Therefore, both PCa and BPH are threats to the health of older men. In recent decades, because of changes in the global population structure and aging, elderly males have increased in number[23, 24]. Thus, the incidence of PCa and BPH will become more common. Understanding the relationship between PCa and BPH may help better predict the occurrence of PCa and may relieve pressure on the medical system. Although PCa primarily arises in the epithelial cells in the peripheral zone, and BPH arises in the transition zone, PCa and BPH have important factors in common, such as growth depend on hormone and responsiveness to antiandrogen therapy[25]. Moreover, inflammation could be an underlying cause of both BPH and PCa. In a study of 180 men with suspected PCa who were biopsied at baseline and after 5 years of follow-up, the 5-year PCa incidence was 20% for men with biopsy specimens showing inflammation at baseline compared with 6% for men with no evidence of inflammation in baseline biopsies[26]. The Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial found that patients with BPH all have chronic inflammatory infiltration[27]. These studies all provide evidence of a relationship between BPH and PCa. Sommers first reported the coexistence of BPH and PCa in 1957[2]. They found that BPH was identified in 80% of cadavers with PCa, and 45% without PCa, respectively. In 1974, Armenian et al. found that patients with BPH had a 4 to 5 times increased risk of PCa. A study by Chokkalingam et al. investigated about 87,000 men, and found that patients with BPH had a 1.2 to 1.7 folds increased risk of PCa incidence and mortality[20]. The largest study was carried out on the Danish male population (about 3,000,000 men) from 1980 to 2007; this study indicated that patients with BPH had a 2 to 3 times increased risk of PCa and a 2 to 8 folds risk of cancer-related mortality[6].

Over recent decades, bioinformatics on microarray data have focused broadly on the PCa occurrence depend on bioinformatic analysis and have revealed some of the mechanisms that may lead to PCa. However, until now, there has been no research concentrated on the molecular mechanisms underlying the development from BPH to PCa. We first adopted an integrated bioinformatics approach to directly compare the differences in gene expression between BPH and PCa. Three datasets, GSE5377, GSE104749, and GSE30994, were analyzed and normalized by R language. A total of 60 DEGs were found, 15 that were up-regulated and 45 that were down-regulated. Then, GO and KEGG pathways were utilized to show how the DEGs’ expression affected the biological pathways underlying the conversion of BPH to PCa. DEGs were mainly enriched in pathways related to the regulation of tight junctions, leukocyte transendothelial migration, and vascular smooth muscle contraction. Previous studies have proven that tight junctions can affect the metastasis of PCa cell lines DU145 and PC-3[28]. Leukocyte transendothelial migration can impact the migration of PCa cells[29]. Vascular smooth muscle contraction can also affect PCa development[30].
Hub genes, MYC, CXCR4, CSRP1, SNAI2, MYL9, ACTG2, and MYH11, were found by analyzing the PPIs of the DEGs. This result indicates that mutations in these genes may play significant roles in the development of PCa from BPH. MYC (MYC proto-oncogene) affects PCa progression due to a high-fat diet and plays a positive role in regulating the androgen receptor and androgen-receptor splice variants in PCa[31, 32]. In addition to PCa, MYC is also correlated with the occurrence of multiple tumors. Furthermore, MYC activation can promote bladder cancer proliferation and invasion[33]. The higher expression of MYC was found in colorectal cancer patients, and it may promote tumorigenesis in colorectal cancer through regulation of the Wnt-and Ras-dependent signal transduction pathway[34, 35]. MYC regulated by BRD4 can promote gastric cancer progression[36]. CXCR4 (C-X-C motif chemokine receptor 4) may promote PCa metastasis through regulation of phosphatidylinositol 4-kinase IIIα (PI4KIIIα) and SAC1 phosphatase[37]. In addition, CXCL12/CXCR4 can increase the malignancy of breast cancer and cervical cancer[38, 39]. SNAI2 (snail family transcriptional repressor 2) can regulate prostate tumor progression and angiogenesis, and metastasis potentially by modulating the GSK-3β/β-catenin signaling pathway[30]. Moreover, research has shown that SNAI2 can promote the invasion and migration of clear cell renal cell carcinoma. Du et al. found that transactivation of SNAI2 and c-MET could promote colorectal cancer metastasis[40]. SNAI2 also participates in regulation of the initiation and metastasis of breast cancer cells[41]. MYL9 (myosin light chain 9) can predict malignant progression and poor biochemical recurrence-free survival of PCa[42]. MYL9 is also associated with the recurrence of colorectal cancer[43]. ACTG2 (actin gamma 2, smooth muscle) and MYH11 (myosin heavy chain 11) have an affection in the development of PCa[44, 45]. ACTG2 can also affect hepatocellular carcinoma cell migration and tumor metastasis, and MYH11 may play a pivotal function in the progression of lung cancer and bladder cancer[46, 47].

When analyzing hub gene expression, 2 up-regulated hub genes, MYC and CXCR4, were not expressed differently between normal prostate samples and PCa samples in the GEPIA database. These results may due to that the normal sequence from GEPIA includes data from both TCGA and GTEx database. Therefore, an additional analysis was conducted; we analyzed hub genes’ expression in tumor and normal samples in UALCAN which include the sequence data only from TCGA database. We found that CXCR4 was not expressed differently in normal and PCa samples in UALCAN. We analyzed the value of hub genes in diagnosing PCa and tumor staging. According to the results, we found that all the hub genes, except CXCR4, could be used as diagnostic markers for PCa progression. In addition, some hub gene can also change when PCa progress. That means these hub genes can be indicators to predict disease progression. In addition, the hub genes were expressed differently in BPH and PCa samples, indicating these genes are potential predictors of PCa development from BPH. Some studies have reported that patients who develop PCa from BPH would have higher cancer-related mortalities. Therefore, at the same time, we also carried out a survival analysis to determine whether the expression of these hub genes affect patients’ survival times. We found that the hub genes do not impact overall survival time, but they can affect progression-free survival time.

In addition, we constructed regression model to value the mutation risk of hub genes in leading to PCa. In the model, we found that all the hub genes may be risk factors in causing PCa except CXCR4. In
nomogram, we found that the mutation of \textit{CSRP1} is the highest risk factor in leading to PCa among all 7 hub genes. Furthermore, the hub genes function in tumor progression also be analyzed. Some hub genes such as \textit{CSRP1}, \textit{MYL9} and \textit{SNAI2} will change when PCa progressed. These results reflect that the expression level change of these hub genes can be potential signal of disease progression.

Finally, we utilized our clinical specimens and C4-2 PCa cells to validate the hub genes’ functions in promoting PCa development. We chose \textit{MYC}, \textit{MYL9} and \textit{SNAI2}. The same results were found in this experiment. The expression of MYC, MYL9 and SNAI2 were obviously changed at the mRNA and protein levels. IHC analysis verified this result. Experiments from C4-2 prostate cancer cell reflects that the 3 hub genes can also affect cell invasion and proliferation ability.

However, our study had some limitations. First, although we included 3 datasets in the study, the number of samples was still small; there were only 10 BPH samples and 24 PCa samples. A too-small sample size may have led to a less representative study. Second, although we validated the hub genes’ expression with TCGA database and experiments, our results may be biased. Thus, our results require further validation. However, our study is the first one to address genes differentially expressed in BPH and PCa by bioinformatic analysis.

\textbf{Conclusion}

We used bioinformatic analysis to identify significant genes in the development of PCa from BPH and validated their roles in PCa. The discovery of these genes may help us better understand the process of BPH conversion to PCa. Meanwhile, BPH patients with these DEGs may have a higher probability of PCa occurrence; therefore, these genes are potential predictors of PCa.

\textbf{Abbreviations}

BPH   benign prostate hyperplasia

PCa   prostate cancer

GEO   Gene Expression Omnibus

GO    Gene oncology

KEGG  Kyoto Encyclopedia of Genes and Genomes

PPI   Protein-protein interaction

TCGA  The Cancer Genome Atlas

GEPIA Gene Expression Profiling Interactive Analysis Platform

ROC   receiver operating characteristic
AUC  area under curve
DCA  decision curve analysis
OS   Overall survival
DFS  Disease-free survival
TNM  Tumor-Node-Metastasis
MYL9  myosin light chain 9
SNAI2  snail family transcriptional repressor 2

Declarations

Acknowledgments

Not applicable

Authors' contributions

Xi Chen and Junjie Ma put forward the idea of the article, wrote the manuscript and analyzed the data. Licheng Wang and Yicong Yao collected the data from GEO and TCGA database. Chengdang Xu and Xinan Wang finished the RT-qPCR, Western blot and IHC experiments. Xi Chen and Tong Zi finished cell invasion and proliferation experiments. Cuidong Bian and Denglong Wu help collecting clinical specimens. Denglong Wu and Gang Wu supervised the experiments progress and revised the manuscript.

Funding

This study was supported by National Natural Science Foundation of China (81672526, 81802560) and Clinical Research Plan of SHDC (NO. SHDC2020CR3074B) and New Frontier Technology Joint Research Project of Shanghai Municipal Hospital (No. SHDC12019112) and Clinical project of Shanghai Municipal Health Commission (NO.20184Y0263, NO.2018Y0105).

Availability of data and materials

The datasets supporting the conclusions of this article are available in the GEO (http://www.ncbi.nlm.nih.gov/geo/), TCGA-PRAD (https://portal.gdc.cancer.gov/) and The Human Protein Atlas (https://www.proteinatlas.org/) database.

Ethic approval and consent participate

The study was approved by the ethic committee of Tongji Hospital, School of Medicine, Tongji University. Each participate volunteered to join and signed the informed consent form.
Consent for publication

Not applicable

Competing interests

All authors declare that they have no competing interests

Contributor Information

Denglong Wu, Email: wudenglong2009@tongji.edu.cn

Gang Wu, Email: wu_urologist@163.com

References

1. Orsted DD, Bojesen SE. The link between benign prostatic hyperplasia and prostate cancer. Nat Rev Urol. 2013;10:49–54.
2. Sommers SC. Endocrine changes with prostatic carcinoma. Cancer. 1957;10:345–58.
3. Chokkalingam AP, Nyren O, Johansson JE, Gridley G, McLaughlin JK, Adami HO, Hsing AW. Prostate carcinoma risk subsequent to diagnosis of benign prostatic hyperplasia: a population-based cohort study in Sweden. Cancer. 2003;98:1727–34.
4. Hammarsten J, Hogstedt B. Calculated fast-growing benign prostatic hyperplasia—a risk factor for developing clinical prostate cancer. Scand J Urol Nephrol. 2002;36:330–8.
5. Briganti A, Chun FK, Suardi N, Gallina A, Walz J, Graefen M, Shariat S, Ebersdobler A, Rigatti P, Perrotte P, Saad F, Montorsi F, Huland H, Karakiewicz PI. Prostate volume and adverse prostate cancer features: fact not artifact. Eur J Cancer. 2007;43:2669–77.
6. Orsted DD, Bojesen SE, Nielsen SF, Nordestgaard BG. Association of clinical benign prostate hyperplasia with prostate cancer incidence and mortality revisited: a nationwide cohort study of 3,009,258 men. Eur Urol. 2011;60:691–8.
7. Luo J, Duggan DJ, Chen Y, Sauvageot J, Ewing CM, Bittner ML, Trent JM, Isaacs WB. Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. Cancer Res. 2001;61:4683–8.
8. Muller I, Urban K, Pantel K, Schwarzenbach H. Comparison of genetic alterations detected in circulating microsatellite DNA in blood plasma samples of patients with prostate cancer and benign prostatic hyperplasia. Ann N Y Acad Sci. 2006;1075:222–9.
9. Shah US, Getzenberg RH. Fingerprinting the diseased prostate: associations between BPH and prostate cancer. J Cell Biochem. 2004;91:161–9.
10. Guo L, Lin M, Cheng Z, Chen Y, Huang Y, Xu K. Identification of key genes and multiple molecular pathways of metastatic process in prostate cancer. PeerJ. 2019;7:e7899.
11. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets – update. Nucleic Acids Res. 2013;41:D991–5.

12. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy – analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20:307–15.

13. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4:P3.

14. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3:Article3.

15. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45:D362–8.

16. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N. Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

17. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45:W98–102.

18. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-Analyzed Tumors. Cell 2018; 173: 530.

19. Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW, Amin MB. Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. Eur Urol. 2018;73:560–9.

20. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varambally S. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia. 2017;19:649–58.

21. Basha O, Shpringer R, Argov CM, Yeger-Lotem E. The DifferentialNet database of differential protein-protein interactions in human tissues. Nucleic Acids Res. 2018;46:D522–6.

22. Alcaraz A, Hammerer P, Tubaro A, Schroder FH, Castro R. Is there evidence of a relationship between benign prostatic hyperplasia and prostate cancer? Findings of a literature review. Eur Urol. 2009;55:864–73.

23. Bushman W. Etiology, epidemiology, and natural history of benign prostatic hyperplasia. Urol Clin North Am. 2009;36:403–15, v.

24. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69:7–34.

25. Andriole GL, Bostwick DG, Brawley OW, Gomella LG, Marberger M, Montorsi F, Pettaway CA, Tammela TL, Teloken C, Tindall DJ, Somerville MC, Wilson TH, Fowler IL. Rittmaster RS and Group RS. Effect of dutasteride on the risk of prostate cancer. N Engl J Med. 2010;362:1192–202.

26. MacLennan GT, Eisenberg R, Fleshman RL, Taylor JM, Fu P, Resnick MI, Gupta S. The influence of chronic inflammation in prostatic carcinogenesis: a 5-year followup study. J Urol. 2006;176:1012–6.
27. Nickel JC, Roehrborn CG, O'Leary MP, Bostwick DG, Somerville MC, Rittmaster RS. The relationship between prostate inflammation and lower urinary tract symptoms: examination of baseline data from the REDUCE trial. Eur Urol. 2008;54:1379–84.

28. Gao F, Alwhaibi A, Artham S, Verma A, Somanath PR. Endothelial Akt1 loss promotes prostate cancer metastasis via beta-catenin-regulated tight-junction protein turnover. Br J Cancer. 2018;118:1464–75.

29. Drake JM, Strohbehn G, Bair TB, Moreland JG, Henry MD. ZEB1 enhances transendothelial migration and represses the epithelial phenotype of prostate cancer cells. Mol Biol Cell. 2009;20:2207–17.

30. Tian X, Tao F, Zhang B, Dong JT, Zhang Z. The miR-203/SNAI2 axis regulates prostate tumor growth, migration, angiogenesis and stemness potentially by modulating GSK-3beta/beta-CATENIN signal pathway. IUBMB Life. 2018;70:224–36.

31. Bai S, Cao S, Jin L, Kobelski M, Schouest B, Wang X, Ungerleider N, Baddoo M, Zhang W, Corey E, Vessella RL, Dong X, Zhang K, Yu X, Flemington EK, Dong Y. A positive role of c-Myc in regulating androgen receptor and its splice variants in prostate cancer. Oncogene. 2019;38:4977–89.

32. Labbe DP, Zadra G, Yang M, Reyes JM, Lin CY, Cacciatore S, Ebot EM, Creech AL, Giunchi F, Fiorentino M, Elfandy H, Syamala S, Karoly ED, Alshalalfa M, Erho N, Ross A, Schaeffer EM, Gibb EA, Takhar M, Den RB, Lehrer J, Karnes RJ, Freedland SJ, Davicioni E, Spratt DE, Ellis L, Jaffe JD, D'Amico AV, Kantoff PW, Bradner JE, Mucci LA, Chavarro JE, Loda M, Brown M. High-fat diet fuels prostate cancer progression by rewiring the metabolome and amplifying the MYC program. Nat Commun. 2019;10:4358.

33. Zhuang C, Ma Q, Zhuang C, Ye J, Zhang F, Gui Y. LncRNA GClnc1 promotes proliferation and invasion of bladder cancer through activation of MYC. FASEB J. 2019;33:11045–59.

34. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487:330–7.

35. Sikora K, Chan S, Evan G, Gabra H, Markham N, Stewart J, Watson. J c-myc oncogene expression in colorectal cancer Cancer. 1987;59:1289–95.

36. Ba M, Long H, Yan Z, Wang S, Wu Y, Tu Y, Gong Y, Cui S. BRD4 promotes gastric cancer progression through the transcriptional and epigenetic regulation of c-MYC. J Cell Biochem. 2018;119:973–82.

37. Sbrissa D, Semaan L, Govindarajan B, Li Y, Caruthers NJ, Stemmer PM, Cher ML, Sethi S, Vaishampayan U. Shisheva A and Chinni SR. A novel cross-talk between CXCR4 and PI4KIIIalpha in prostate cancer cells. Oncogene. 2019;38:332–44.

38. Lecavalier-Barsoum M, Chaudary N, Han K, Koritzinsky M, Hill R, Milosevic M. Targeting the CXCL12/CXCR4 pathway and myeloid cells to improve radiation treatment of locally advanced cervical cancer. Int J Cancer. 2018;143:1017–28.

39. Martinez-Ordonez A, Seoane S, Cabezas P, Eiro N, Sendon-Lago J, Macia M, Garcia-Caballero T, Gonzalez LO, Sanchez L, Vizoso F, Perez-Fernandez R. Breast cancer metastasis to liver and lung is facilitated by Pit-1-CXCL12-CXCR4 axis. Oncogene. 2018;37:1430–44.

40. Du F, Li X, Feng W, Qiao C, Chen J, Jiang M, Qiu Z, Qian M, Tian D, Nie Y, Fan D, Wu K. and Xia L. SOX13 promotes colorectal cancer metastasis by transactivating SNAI2 and c-MET. Oncogene.
2020;39:3522–40.

41. Jiang B, Chen W, Qin H, Diao W, Li B, Cao W, Zhang Z, Qi W, Gao J, Chen M, Zhao X, Guo H. TOX3 inhibits cancer cell migration and invasion via transcriptional regulation of SNAI1 and SNAI2 in clear cell renal cell carcinoma. Cancer Lett. 2019;449:76–86.

42. Huang YQ, Han ZD, Liang YX, Lin ZY, Ling XH, Fu X, Cai C, Bi XC, Dai QS, Chen JH, He HC, Chen YR, Jiang FN, Zhong WD. Decreased expression of myosin light chain MYL9 in stroma predicts malignant progression and poor biochemical recurrence-free survival in prostate cancer. Med Oncol. 2014;31:820.

43. Qiu X, Cheng SH, Xu F, Yin JW, Wang LY, Zhang XY. Weighted gene co-expression network analysis identified MYL9 and CNN1 are associated with recurrence in colorectal cancer. J Cancer. 2020;11:2348–59.

44. Lai J, Moya L, An J, Hoffman A, Srinivasan S, Panchadsaram J, Walpole C, Perry-Keene JL, Chambers S, Australian Prostate Cancer B, Lehman ML, Nelson CC, Clements JA, Batra J. A microsatellite repeat in PCA3 long non-coding RNA is associated with prostate cancer risk and aggressiveness. Sci Rep. 2017;7:16862.

45. Wu Y, Liu ZG, Shi MQ, Yu HZ, Jiang XY, Yang AH, Fu XS, Xu Y, Yang S, Ni H, Shen SJ, Li WD. Identification of ACTG2 functions as a promoter gene in hepatocellular carcinoma cells migration and tumor metastasis. Biochem Biophys Res Commun. 2017;491:537–44.

46. Hu J, Zhou L, Song Z, Xiong M, Zhang Y, Yang Y, Chen K, Chen Z. The identification of new biomarkers for bladder cancer: A study based on TCGA and GEO datasets. J Cell Physiol 2019.

47. Nie MJ, Pan XT, Tao HY, Xu MJ, Liu SL, Sun W, Wu J, Zou X. Clinical and prognostic significance of MYH11 in lung cancer. Oncol Lett. 2020;19:3899–906.

Figures
Figure 1

Venn map reflects different expressed genes (DEGs) between benign prostate hyperplasia (BPH) and prostate cancer (PCa) from GSE5377, GSE104749 and GSE30994. (A) The up-regulated DEGs from GSE5377, GSE104749 and GSE30994. (B) The down-regulated DEGs from GSE5377, GSE104749 and GSE30994.
Figure 2

Protein-protein interaction (PPI) network of DEGs was shown (Red color represents up-regulated DEGs, Blue color represents down-regulated DEGs); GO and KEGG pathways was shown in scatter map. (A) The PPI network of all DEGs. (B) The interaction of 7 hub genes in PPI network. (C) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of DEGs was shown in scatter map. (D) The Gene Oncology (GO) analysis of DEGs was exhibited in scatter map.
Figure 3

The expression of hub genes between normal prostate tissues and prostate cancer samples from GEPIA depend on TCGA and GTEx database. (A) MYC (B) CXCR4 (C) CSRP1 (D) SNAI2 (E) MYL9 (F) ACTG2 (G) MYH11 Red color represents tumor samples and gray color represents normal samples. * represent statistical differences between two groups.
Figure 4

The diagnostic model of hub genes was constructed. (A) The Receiver Operating Characteristic (ROC) curve of hub genes was constructed. (B) The Decision Curve Analysis (DCA) of hub genes was shown.
Figure 5

The expression of hub genes in different tumor stages depend on TCGA database. (A) The expression of hub genes in different Tumor (T) stage depend on TNM classification of malignant tumors. (B) The expression of hub genes in different Node (N) stage depend on TNM classification of malignant tumors. (C) The expression of hub genes in different T stage depend on ACJJ tumor classification.
Figure 6

Risk prediction model of hub genes mutation in leading to PCa by logistic regression depend on TCGA data. (A) The nomogram of hub genes expression changes in causing PCa. (B) Calibration plot of actual risk probability and nomogram risk of nomogram. (C) The forest map of single factor of regression analysis of hub genes in causing PCa. (D) The forest map of multiple factors of regression analysis of hub genes in leading to PCa.
Figure 7

The disease-free survival (DFS) of 7 hub genes in patients with PCa was evaluated by Kaplan-Meier curve with higher and lower expression than average expression level from GEPIA. (A) MYC (B) CXCR4 (C) CSRP1 (D) SNAI2 (E) MYL9 (F) ACTG2 (G) MYH11
Figure 8

The expression of hub genes in patients with PCa depend on different database. (A) The expression of 6 hub genes in PCa depend on UALCAN database. (B) The expression of 6 hub genes in PCa depend on The Human Protein Atlas.

Figure 9
The expression of MYC, MYL9 and SNAI2 in patients with PCa from clinical specimens. (A-B) The mRNA and protein expression level of MYC in 6 paired patients’ samples from cancer tissues and paracancerous normal tissues. (C-D) The expression of MYL9 at mRNA and protein level in 6 paired patients’ samples from cancer tissues and paracancerous normal tissues. (E-F) The SNAI2 mRNA and protein expression in 6 paired patients from both tumor tissues and paracancerous normal tissues. (G-I) The Immunohistochemical results of MYC, MYL9 and SNAI2 in 2 paired patients’ specimens from both cancer tissues and paracancerous normal tissues. The expression level of mRNA and protein was used GAPDH as inner control. * represent P<0.05, ** represent P<0.01, *** represent P<0.001.
MYC, MYL9 and SNAI2 can affect PCa cell invasion and proliferation. (A-B) The mRNA and protein expression of MYC after shMYC plasmid transfected into C4-2 prostate cancer cell. (C-D) The MYL9 mRNA and protein level after overexpression MYL9 plasmid transfected into C4-2 prostate cancer cell. (E-F) The expression of SNAI2 in both mRNA and protein level after C4-2 cell transfected by overexpression SNAI2 plasmid. (G) The invasion ability of C4-2 cells after transfection with different plasmid and normal
control. (H) The proliferation ability of C4-2 cells after transfection with different plasmid and normal control. * represent $P<0.05$, ** represent $P<0.01$, *** represent $P<0.001$.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1.tif
- FigS2.tif