

Identification of *PTPRR* and *JAG1* as key genes in castration-resistant prostate cancer by integrated bioinformatics methods[#]

Ji-li WANG^{1,2}, Yan WANG^{1,2}, Guo-ping REN^{†‡1,2}

¹Department of Pathology, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China

²Department of Pathology and Pathophysiology, School of Medicine, Zhejiang University, Hangzhou 310058, China

[†]E-mail: gpren2002@163.com

Received June 20, 2019; Revision accepted Oct. 22, 2019; Crosschecked Feb. 3, 2020

Abstract: To identify novel genes in castration-resistant prostate cancer (CRPC), we downloaded three microarray datasets containing CRPC and primary prostate cancer in Gene Expression Omnibus (GEO). R packages *affy* and *limma* were performed to identify differentially expressed genes (DEGs) between primary prostate cancer and CRPC. After that, we performed functional enrichment analysis including gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway. In addition, protein-protein interaction (PPI) analysis was used to search for hub genes. Finally, to validate the significance of these genes, we performed survival analysis. As a result, we identified 53 upregulated genes and 58 downregulated genes that changed in at least two datasets. Functional enrichment analysis showed significant changes in the positive regulation of osteoblast differentiation pathway and aldosterone-regulated sodium reabsorption pathway. PPI network identified hub genes like cortactin-binding protein 2 (*CTTNBP2*), Rho family guanosine triphosphatase (GTPase) 3 (*RND3*), protein tyrosine phosphatase receptor-type R (*PTPRR*), Jagged1 (*JAG1*), and lumican (*LUM*). Based on PPI network analysis and functional enrichment analysis, we identified two genes (*PTPRR* and *JAG1*) as key genes. Further survival analysis indicated a relationship between high expression of the two genes and poor prognosis of prostate cancer. In conclusion, *PTPRR* and *JAG1* are key genes in the CRPC, which may serve as promising biomarkers of diagnosis and prognosis of CRPC.

Key words: Bioinformatics; Protein tyrosine phosphatase receptor-type R (*PTPRR*); Jagged1 (*JAG1*); Differentially expressed genes (DEGs); Castration-resistant prostate cancer (CRPC); Functional enrichment
<https://doi.org/10.1631/jzus.B1900329>

CLC number: R737.25

1 Introduction


Prostate cancer (PCa) is one of the most commonly diagnosed cancers in the world. According to Cancer Statistics 2019, PCa became the first among estimated new cases of cancer in men and the second among estimated reasons behind deaths in men (Siegel et al., 2019). Similarly, the incidence

rates and mortality rates of PCa have been continuously rising in China (Chen et al., 2016). Thus, PCa has received considerable critical attention. One of the main challenges of PCa is its treatment, especially the treatment of castration-resistant PCa (CRPC). CRPC has poor prognosis due to its aggressive behavior. It has been found that even PCa patients who initially responded well to the androgen deprivation therapy, they will eventually develop CRPC and the outcomes are not satisfactory (Heidenreich et al., 2014). Thus, a key issue in PCa is the identification of biomarkers and treatment targets for CRPC.

Recently, many mechanisms of castration resistance have been proposed, including de novo or

[‡] Corresponding author

[#] Electronic supplementary materials: The online version of this article (<https://doi.org/10.1631/jzus.B1900329>) contains supplementary materials, which are available to authorized users

 ORCID: Ji-li WANG, <https://orcid.org/0000-0002-8976-508X>

© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2020

cytochrome P450 family-induced synthesis of androgen, upregulation of androgen metabolism-related genes, and dysregulation of tumor microenvironment (Watson et al., 2015). Despite acknowledgement of these mechanisms, treatment of CRPC remains a challenge (Heidenreich et al., 2014). Thus, a better understanding of CRPC and identification of its biomarkers or targets, which may open a new window of opportunity for combating the disease, are urgently needed (Attard et al., 2016). Previous studies have identified numerous biomarkers and targets in CRPC by microarrays. For example, D'Antonio et al. (2008) compared androgen-independent PCa cells with androgen-dependent LNCaP cells, to identify pathways to androgen independence. Terada et al. (2010) compared gene expression between a mouse xenograft model of PCa with its later CRPC and identified prostaglandin E2 receptor subtype 4 (*EP4*) as a central gene of CRPC. These microarray data led us to think the possibility of using combined datasets and integrated bioinformatics to explore the molecular and cellular characteristics of CRPC. To our knowledge, there are few studies that have used bioinformatics to identify gene alterations between primary PCa and CRPC.

In this study, we used three microarray datasets in Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo>) with the R packages *affy* and *limma* to identify differentially expressed genes (DEGs) between CRPC and primary PCa. Furthermore, functional enrichment analysis, protein-protein interaction (PPI) network, and survival analysis provided us with more information about CRPC. In brief, we identified protein tyrosine phosphatase receptor-type R (*PTPRR*) and *Jagged1* (*JAG1*) as key genes in CRPC, which may serve as promising biomarkers for diagnosis and prognosis, even as therapeutic targets of CRPC.

2 Materials and methods

2.1 Microarray data

Three gene expression profiles (GSE8702, GSE21887, and GSE33316) were obtained from the GEO. These gene expression profiles include cell line (LNCaP from GSE8702) (D'Antonio et al., 2008) and xenograft models (KUCaP-2 from GSE21887 and LuCaP35 from GSE33316) (Terada et al., 2010; Sun et al., 2012), which could mimic CRPC (Table 1). In addition, the platform of the three datasets was GPL570, which means that these microarrays detected the same gene pools, providing the comparability of these datasets.

2.2 Identification of DEGs

R software (Version 3.5.1; <https://www.r-project.org>) was used for the identification of DEGs. We downloaded the raw data from GEO (GSE8702, GSE21887, and GSE33316), normalized them, and transformed them into expression values by the *affy* package of Bioconductor (<http://www.bioconductor.org>) (Gautier et al., 2004). Then we used empirical Bayes methods to identify DEGs between CRPC and primary androgen-dependent PCa with *limma* package (Ritchie et al., 2015). The criteria of DEGs were false discovery rate (FDR) < 0.05 and $|\log_2(\text{fold change})| > 1$ (Su et al., 2018). Then we used Venn diagrams to find the common altered genes (including increased and decreased genes) in the three datasets.

2.3 Functional enrichment analysis of the DEGs

The Database for Annotation, Visualization, and Integrated Discovery (DAVID; <http://www.david.niaid.nih.gov>) (Dennis et al., 2003) is a useful tool for functional annotation of DEGs via four web-based analysis modules including gene ontology (GO) charts (Gene Ontology Consortium, 2004) and Kyoto

Table 1 Characteristics of the microarray datasets from GEO database

Accession/ID	PMID	CRPC group	Control group
GSE8702	18302219 (D'Antonio et al., 2008)	Androgen deprived for 11 or 12 months	Androgen deprived for three weeks or one month
GSE21887	20145136 (Terada et al., 2010)	Castration resistance in xenograft mice	Androgen-dependent xenograft mice
GSE33316	22108827 (Sun et al., 2012)	Androgen-deprived LuCaP35 xenografts in castrated mice	LuCaP35 xenografts from non-treated mice

GEO: Gene Expression Omnibus; PMID: PubMed unique identifier; CRPC: castration-resistant prostate cancer

encyclopedia of genes and genomes (KEGG) charts (Kanehisa and Goto, 2000; Zhu et al., 2019). We used the DAVID database to perform GO and KEGG pathway analyses on DEGs that were upregulated or downregulated in at least two datasets.

2.4 PPI network construction and analysis

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; <https://string-db.org>) is used to construct PPI network, which shows physical and functional interactions (Szklarczyk et al., 2017; Lin et al., 2018). In this study, the protein pairs with combined scores of >0.15 were selected for the PPI network construction. Further, the Cytoscape software (Version 3.6.1) (Shannon et al., 2003) was utilized to calculate the node degree by Network Analyzer app and draw the PPI network with different colors and sizes, which present the regulation (up or down) and node degree, respectively.

2.5 Expression validation and survival analysis of the key genes

Expression validation was performed in another GEO dataset (GSE70770-GPL10558), which contains 13 patients with CRPC and 206 patients with primary

PCa (the clinical and pathological characteristics were shown in Table S1). Survival analysis was performed in Gene Expression Profiling Interactive Analysis (GEPIA; <http://gepia.cancer-pku.cn/index.html>) (Tang et al., 2017), an interactive web application for gene expression and survival analysis based on The Cancer Genome Atlas (TCGA; <http://www.cancergenome.nih.gov>) database (Weinstein et al., 2013), which offers abundant clinical information from a huge PCa sample size (Guo et al., 2018).

3 Results

3.1 DEG identification

After performing DEG analysis with R package, we identified the DEGs in the CRPC samples compared with the primary PCa samples. The GSE8702 dataset harbors 616 upregulated genes and 1828 downregulated genes. The GSE21887 dataset identified 481 increased genes and 25 decreased genes, while the GSE33316 dataset found 149 genes with enhanced expression and 534 genes with suppressed expression. These DEGs are shown in volcano maps and heat maps (Figs. 1 and 2).

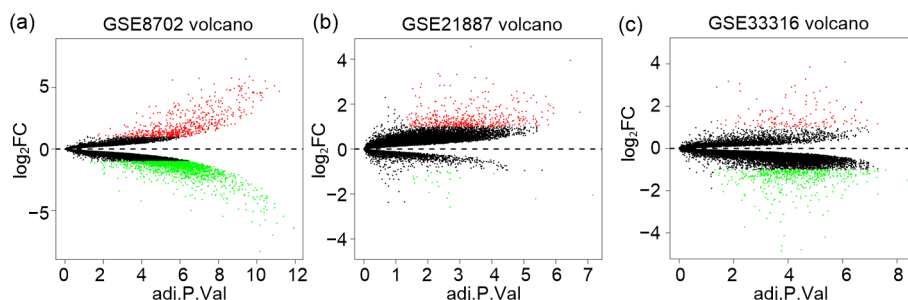


Fig. 1 Volcano plots of all genes in GSE8702, GSE21887, and GSE33316

Red dots represent genes with fold change ≥ 2 and $P < 0.05$, green dots represent genes with fold change ≤ -2 and $P < 0.05$, and the other dots represent the rest of genes with no statistically significant change in expression. FC, fold change

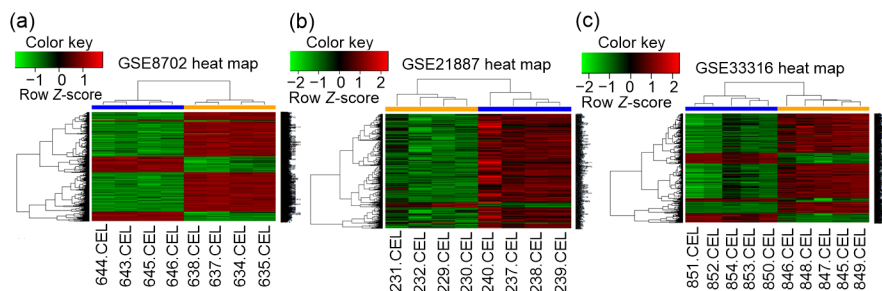


Fig. 2 Heat maps of differential expressed genes (DEGs) in GSE8702, GSE21887, and GSE33316

Red color represents upregulated genes (fold change ≥ 2 and $P < 0.05$) and green color represents downregulated genes (fold change ≤ -2 and $P < 0.05$)

Then, we performed Venn diagrams, which extracted 53 upregulated genes and 58 downregulated genes that changed in at least two datasets (Fig. 3 and Table 2). In addition, the expression of Rho family guanosine triphosphatase (GTPase) 3 (*RND3*), neuroligin 4 X-linked (*NLGN4X*), and cyclic adenosine monophosphate (cAMP)-dependent protein kinase inhibitor α (*PKIA*) was enhanced, while only ephrin-B2 (*EFNB2*) was decreased in all datasets.

3.2 GO and KEGG pathway enrichment

To explore the functional enrichment of the identified DEGs, we submitted the upregulated genes and downregulated genes, as described in Table 2, to the DAVID for functional pathway analysis including GO and KEGG, and the results are presented in Table 3.

The most significantly altered pathway was positive regulation of osteoblast differentiation pathway. The other significant pathways included: in utero embryonic development, negative regulation of stem

cell differentiation, nucleosome assembly, regulation of G2/M transition of mitotic cell cycle, negative regulation of DNA-templated transcription (Table 3). In addition, the KEGG pathway analysis indicated that the aldosterone-regulated sodium reabsorption played an essential role in CRPC. Regarding molecular function, symporter activity and protein homodimerization activity were essential.

3.3 Construction of PPI network

To explore the protein-protein interaction of the DEGs, we used STRING, a web-based analysis of protein interaction. Under the criteria as described in methods, a total of 82 DEGs among 111 commonly changed DEGs were filtered into the PPI network. The PPI network contained 82 nodes and 201 edges (Fig. 4), which was further analyzed in Cytoscape, with Network Analyzer app for calculating the node degree. The genes cortactin-binding protein 2 (*CTTNBP2*), *RND3*, lumican (*LUM*), Toll-like receptor 3 (*TLR3*), zinc finger protein 423 (*ZNF423*), Rho-related BTB domain containing protein 1 (*RhoBTB1*), leucine-rich repeat and immunoglobulin domain-containing protein 1 (*LINGO1*), trophoblast glycoprotein (*TPBG*), *JAG1*, *PTPRR*, Nance-Horan Syndrome (*NHS*), protein tyrosine phosphatase receptor type β (*PTPRB*), zinc finger and BTB domain containing 16 (*ZBTB16*), heat shock protein family A member 2 (*HSPA2*), cyclin B3 (*CCNB3*), PDZ and LIM domain protein 5 (*PDLIM5*), twist homolog 1 (*TWIST1*), histone cluster 1 H4 family member h (*HIST1H4H*), nuclear receptor subfamily 3 group C member 2 (*NR3C2*), and guanylate cyclase 1 soluble subunit $\alpha 3$ (*GUCY1A3*) were the most significant 20 node degree genes, which were regarded as the hub nodes, since they may be responsible for CRPC progression.

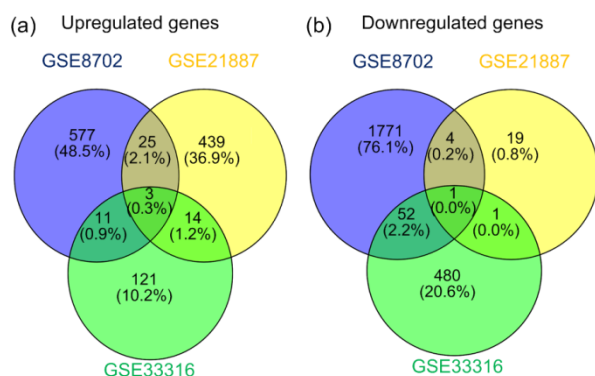


Fig. 3 Venn diagram of differential expressed gene (DEG) selection in three datasets

(a) Venn diagram of upregulated genes (fold change ≥ 2 and $P < 0.05$). (b) Venn diagram of downregulated genes (fold change ≤ -2 and $P < 0.05$)

Table 2 Differential expressed genes (DEGs) in at least two datasets

Alternation	Genes
Upregulated	<i>RND3</i> , <i>NLGN4X</i> , <i>PKIA</i> , <i>FOXQ1</i> , <i>CMPK2</i> , <i>TMEM45A</i> , <i>C1orf168</i> , <i>HLA-DRA</i> , <i>TPBG</i> , <i>MASTL</i> , <i>SPIN4</i> , <i>SPDL1</i> , <i>PSRC1</i> , <i>SKA1</i> , <i>LUM</i> , <i>RFWD3</i> , <i>SACS</i> , <i>WBP5</i> , <i>NUDT11</i> , <i>SLC1A1</i> , <i>ELOVL6</i> , <i>ITM2A</i> , <i>NHS</i> , <i>SSTR5-AS1</i> , <i>PVRL3</i> , <i>CHAC2</i> , <i>RHOBTB1</i> , <i>DNAH5</i> , <i>XDH</i> , <i>PTPRR</i> , <i>PTPRB</i> , <i>JAG1</i> , <i>HAGLROS</i> , <i>TLR3</i> , <i>NBEA</i> , <i>ZNF423</i> , <i>MAP2</i> , <i>SI</i> , <i>AGPAT9</i> , <i>ADD3</i> , <i>HOXA13</i> , <i>DACHI</i> , <i>FAM102B</i> , <i>CLUL1</i> , <i>LINGO1</i> , <i>SCGB1D2</i> , <i>CTTNBP2</i> , <i>RGS2</i> , <i>CYP4X1</i> , <i>APELA</i> , <i>ZBED9</i> , <i>LOC286052</i> , <i>FOXD1</i>
Downregulated	<i>EFNB2</i> , <i>TWIST1</i> , <i>GUCY1A3</i> , <i>CTBP1-AS</i> , <i>ADAMTS1</i> , <i>CMTR2</i> , <i>HSD11B2</i> , <i>ABHD2</i> , <i>FAM110C</i> , <i>SLC25A21</i> , <i>PP14571</i> , <i>CEBPD</i> , <i>SLC2A10</i> , <i>HSPA2</i> , <i>GCNT2</i> , <i>SCNN1B</i> , <i>ACOT9</i> , <i>NPNT</i> , <i>HIST1H2BE</i> , <i>TRIM6</i> , <i>NR3C2</i> , <i>LINC01139</i> , <i>HHEX</i> , <i>CYB5B</i> , <i>KRT20</i> , <i>ERO1LB</i> , <i>HIST1H4H</i> , <i>FZD3</i> , <i>HIST1H1C</i> , <i>ZNF883</i> , <i>PDLIM5</i> , <i>RRN3</i> , <i>CCNB3</i> , <i>PFN2</i> , <i>SLC16A9</i> , <i>VLDLR</i> , <i>ZBTB16</i> , <i>MAN2A1</i> , <i>ADRB2</i> , <i>GCHFR</i> , <i>RBPM52</i> , <i>KCTD9</i> , <i>ELOVL2-AS1</i> , <i>GADD45G</i> , <i>LOC100288675</i> , <i>SLC16A6</i> , <i>AIFM2</i> , <i>ENDOD1</i> , <i>ABAT</i> , <i>ZFYVE21</i> , <i>GBA3</i> , <i>BMP1B</i> , <i>MAK</i> , <i>H2BFS</i> , <i>NANOS1</i> , <i>EOGT</i> , <i>FXD3</i> , <i>CUX2</i>

Table 3 Functional enrichment analysis

Pathway ID	Pathway description	Count	P-value	Genes
GO-BP				
GO:0045669	Positive regulation of osteoblast differentiation	4	0.004615969	<i>CEBPD, NPNT, JAG1, BMPR1B</i>
GO:0001701	In utero embryonic development	5	0.020901560	<i>MAN2A1, HHEX, RRN3, PTPRR, TWIST1</i>
GO:0045892	Negative regulation of DNA-templated transcription	8	0.021870896	<i>HHEX, CEBPD, TRIM6, DACH1, ZBTB16, FOXD1, TWIST1, ZNF423</i>
GO:0006334	Nucleosome assembly	4	0.029188620	<i>H2BFS, HIST1H1C, HIST1H2BE, HIST1H4H</i>
GO:0060087	Relaxation of vascular smooth muscle	2	0.033125732	<i>RGS2, GUCY1A3</i>
GO:0010389	Regulation of G2/M transition of mitotic cell cycle	2	0.049279109	<i>CCNB3, PKIA</i>
GO:2000737	Negative regulation of stem cell differentiation	2	0.049279109	<i>TRIM6, JAG1</i>
GO-CC				
GO:0016324	Apical plasma membrane	7	0.004174384	<i>ADRB2, SI, FZD3, JAG1, SCNN1B, NHS, SLC1A1</i>
GO:0043235	Receptor complex	5	0.004393425	<i>PTPRB, ADRB2, NR3C2, BMPR1B, VLDLR</i>
GO:0005887	Integral component of plasma membrane	14	0.031795756	<i>PTPRB, FXYD3, SLC2A10, EFNB2, TLR3, JAG1, TPBG, ADRB2, SLC16A6, NLGN4X, SCNN1B, BMPR1B, SLC1A1, HLA-DRA</i>
GO:0000940	Condensed chromosome outer kinetochore	2	0.040957971	<i>SPDL1, SKA1</i>
GO-MF				
GO:0015293	Symporter activity	3	0.030268219	<i>SLC16A6, SLC16A9, SLC1A1</i>
GO:0042803	Protein homodimerization activity	9	0.040796847	<i>RBPMS2, XDH, HHEX, ADRB2, CEBPD, NLGN4X, ABAT, ZBTB16, TWIST1</i>
KEGG				
hsa04960	Aldosterone-regulated sodium reabsorption	3	0.019200921	<i>NR3C2, HSD11B2, SCNN1B</i>

GO-BP: gene ontology-biological process; GO-CC: gene ontology-cellular component; GO-MF: gene ontology-molecular function; KEGG: Kyoto encyclopedia of genes and genomes

3.4 Identification of key gene signatures in CRPC and survival analysis

Based on PPI network and GO pathway, we identified two genes (*PTPRR* and *JAG1*) among the top 20 nodes in the PPI network as key genes (Fig. 4). Further, we applied expression validation and survival analysis on the two genes to evaluate their roles in prognosis. Both genes were upregulated in the CRPC compared with primary PCa (Fig. 5), and the high expression was associated with shorter disease-free survival (Fig. 6). Interestingly, the expression of *PTPRR* and *JAG1* did not differ between primary PCa and normal prostate tissues (Fig. 7), which led us to hypothesize that these genes may have a role in the progression of PCa rather than in the initiation of PCa.

4 Discussion

In the present study, we identified two significant genes *JAG1* and *PTPRR*, which participate in

various pathways including positive regulation of osteoblast differentiation, stem cell differentiation, and embryonic development. The expression of these genes is robustly upregulated in CRPC rather than in primary PCa. Owing to the importance of these pathways in the progression and metastasis in PCa, we hypothesized that these two genes (*JAG1* and *PTPRR*) play a critical role in the progression of PCa.

In line with our finding, a number of seminal papers have reported that *JAG1* has an essential role in cancers including PCa (Santagata et al., 2004; Zhu et al., 2013; Li et al., 2014; Su et al., 2017). *JAG1* encodes the protein Jagged1, which is regarded as a major Notch ligand. Previous studies have revealed that Notch signaling pathway is an evolutionarily conserved pathway, which contributes to the regulation of tissue homeostasis and carcinogenesis as well as stem cell function, developmental cell fate decisions, and cellular differentiation (Artavanis-Tsakonas et al., 1999). Consistent with its significant role in the development, the Notch signaling pathway has an active role in cancer development (Espinoza et al.,

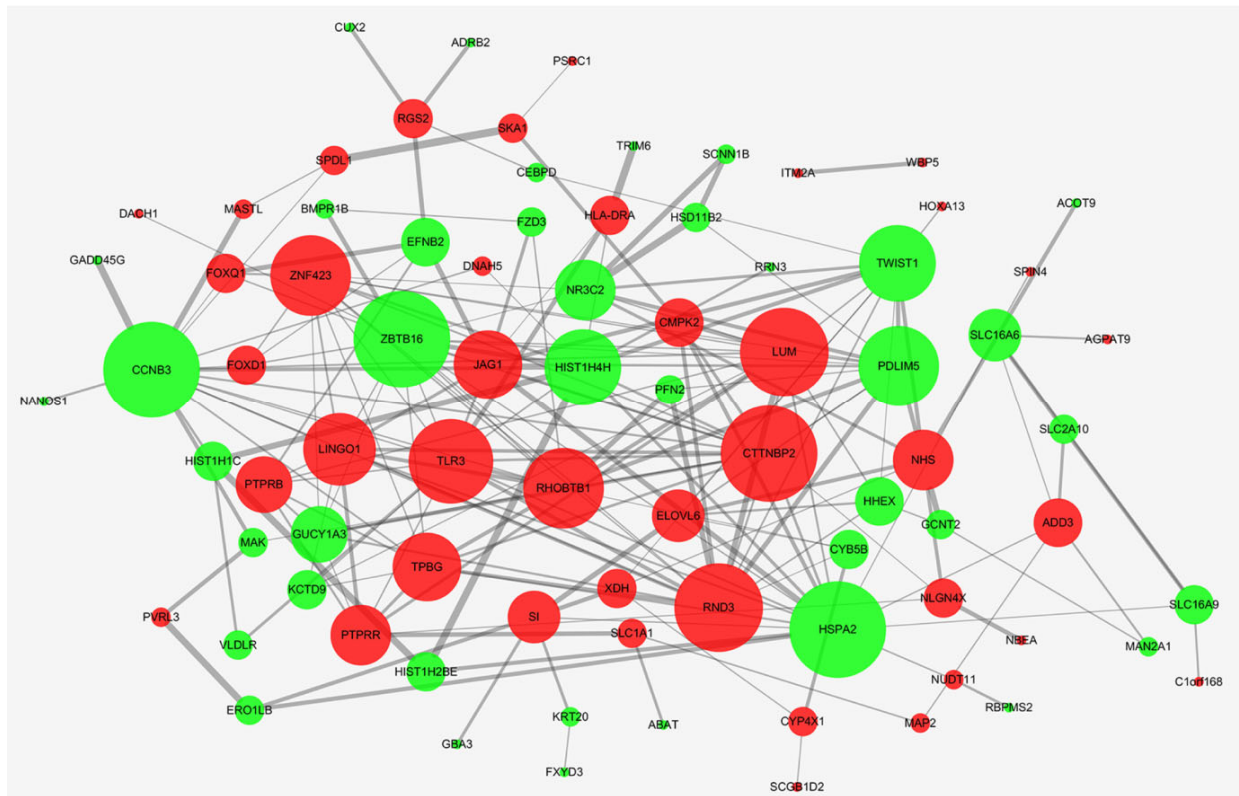


Fig. 4 Construction of protein-protein interaction (PPI) network of differential expressed genes (DEGs)
Red circles are upregulated genes in PPI network, while green ones are downregulated genes

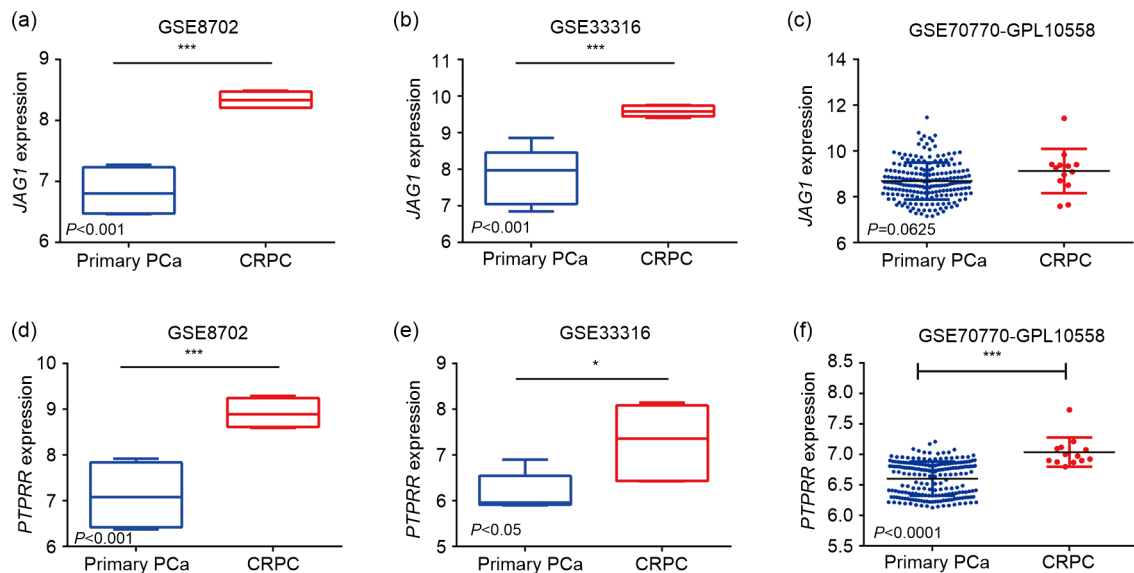


Fig. 5 Expression of key genes (*JAG1* and *PTPRR*) in the GEO datasets

(a–c) Expression of *JAG1* mRNA in GSE8702, GSE33316, and GSE70770-GPL10558, respectively; (d–f) Expression of *PTPRR* mRNA in GSE8702, GSE33316, and GSE70770-GPL10558, respectively. Blue represents primary PCa and red represents CRPC. PCa: prostate cancer; *JAG1*: Jagged1; *PTPRR*: protein tyrosine phosphatase receptor-type R; GEO: Gene Expression Omnibus; CRPC: castration-resistant prostate cancer

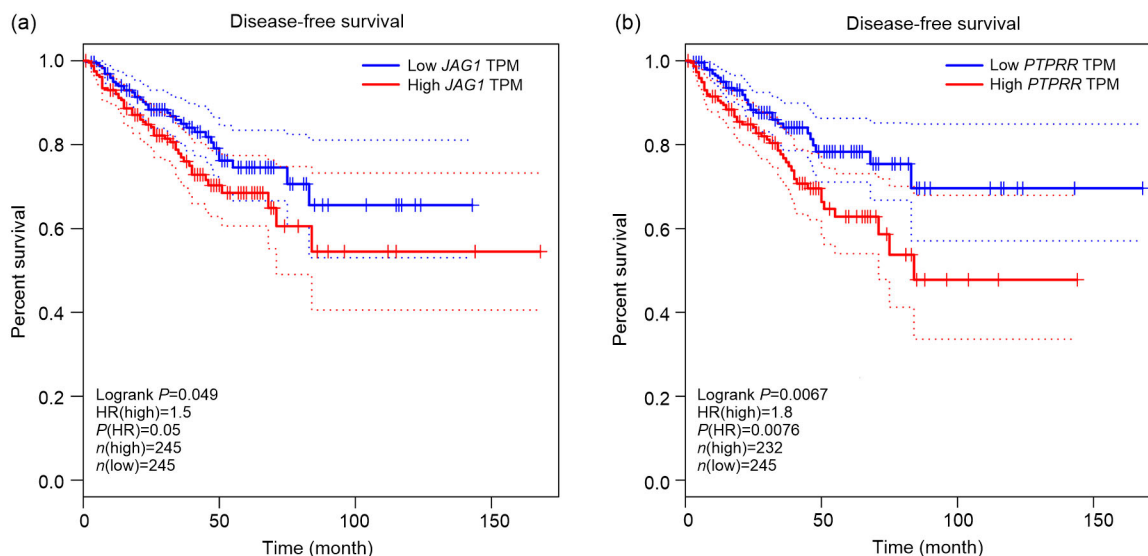


Fig. 6 Disease-free survival in relation to the expression of the key genes (*JAG1* and *PTPRR*) in TCGA (a) *JAG1*, $P=0.049$; (b) *PTPRR*, $P=0.0067$. Red line represents cases with high expression of the genes; blue line represents cases with low expression of the genes. *JAG1*: Jagged1; *PTPRR*: protein tyrosine phosphatase receptor-type R; TCGA: The Cancer Genome Atlas; TPM: transcripts per kilobase of exon model per million mapped reads; HR: hazard ratio

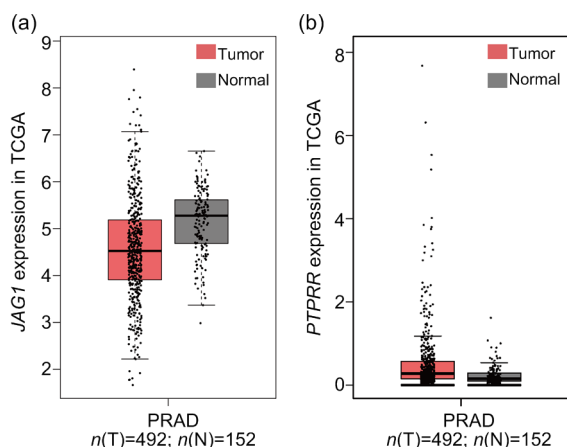


Fig. 7 Expression of the key genes *JAG1* and *PTPRR* in TCGA with GEPIA

(a) *JAG1*; (b) *PTPRR*. Red represents prostate cancer tissues and grey represents normal tissues. *JAG1*: Jagged1; *PTPRR*: protein tyrosine phosphatase receptor-type R; TCGA: The Cancer Genome Atlas; GEPIA: Gene Expression Profiling Interactive Analysis; PRAD: prostate adenocarcinoma

2013). The Notch ligand Jagged1 has also served as a potential therapeutic of cancer (Li et al., 2014). In this study, we found that *JAG1* is upregulated in CRPC, and its high expression is related to poor outcome. This is not the only study demonstrating the important role of *JAG1*. Some reports have shown that *JAG1* is

increased in PCa and is associated with metastasis and recurrence (Santagata et al., 2004; Zhu et al., 2013; Su et al., 2017). Further, it has been shown that inhibition of Notch signaling weakens growth, proliferation, and migration in PCa (Shou et al., 2001; Yong et al., 2011; Kwon et al., 2016), suggesting the critical role of *JAG1* and Notch signaling in PCa. Mechanistically, Jagged1-mediated Notch activation can induce epithelial-mesenchymal transition (EMT), thus contributing to cancer cell motility (Sethi et al., 2011). All in all, *JAG1* and Notch signaling plays a critical role in PCa development. Given the role of *JAG1* and Notch signaling in osteoblast differentiation and the frequent osteoblastic bone metastases (bone forming) in CRPC (Logothetis and Lin, 2005), we hypothesized that *JAG1* may be involved in the bone metastasis of CRPC. Thus, *JAG1* might serve as a therapeutic target for bone metastasis in CRPC.

One unanticipated finding was regarding the gene *PTPRR*, which is considered a tumor suppressor in several cancers. It was found to be overexpressed in CRPC and the high expression was related to poor prognosis. This is extremely striking because *PTPRR* belongs to the family of protein tyrosine phosphatases (PTPs), which regulates proliferation, differentiation, and function of cancer cells by dephosphorylating tyrosine (for example, tyrosine in mitogen-activated

protein kinase (MAPK)) (Noordman et al., 2006; Schmitt et al., 2009). Several studies have shown that the expression of *PTPRR* is decreased in several cancers including colorectal carcinomas, cervical cancer, and oral squamous cell carcinoma due to high methylation (Laczmanska et al., 2013; Su et al., 2013; Chang et al., 2014; Woźniak et al., 2014), suggesting a tumor suppressive role for *PTPRR* (Menigatti et al., 2009; Su et al., 2013; Duś-Szachniewicz et al., 2015). These in vitro studies have also shown that overexpression of *PTPRR* inhibits migration and invasion in these cancers. Further, a study has shown that *PTPRR* is suppressed by androgen in PCa, and negatively regulates the RAS/extracellular signal-regulated protein kinases 1 and 2 (ERK1/2) pathway, thus reducing cell proliferation in response to androgen (Munkley et al., 2015). However, our results strongly indicate that patients with CRPC have higher expression of *PTPRR* and there is a relationship between higher expression of *PTPRR* and poorer outcomes, suggesting a survival advantage for cancer cells overexpressing *PTPRR*. Additionally, we found that the Grasso Prostate dataset also showed that *PTPRR* is upregulated in CRPC (Fig. S1). There are two explanations for this difference. One explanation is that upregulation of *PTPRR* is a mechanism to combat cancers, which means that high expression of *PTPRR* suggests war against cancer cells, and thus the patients may have a poor prognosis. The other explanation is that *PTPRR*, in contrast to other PTP family members, serves as an oncogene. Thus, further study regarding *PTPRR* is necessary to identify the functions of this gene. In summary, *PTPRR* may serve as a prognosis biomarker for PCa.

Although this study is first to identify the two significant genes (*JAG1* and *PTPRR*) in CRPC by integrated bioinformatics analysis, we still need to recognize some limitations. Firstly, the sample we used to identify DEGs was from cell lines and xenograft models, and the sample size was small. Although we used the human sample tissue from another GEO dataset and TCGA to validate our results, the total sample size was still small due to the difficulty of obtaining CRPC sample. Secondly, our results have not been validated by experiments like quantitative real-time polymerase chain reaction (qRT-PCR), western blot, or functional in vitro experiments. Further studies on human tissue samples are needed to validate our results.

5 Conclusions

We identified two essential genes (*JAG1* and *PTPRR*) that contribute to CRPC possibly by regulating key pathways like positive regulation of osteoblast differentiation, stem cell differentiation, and embryonic development. The two genes can be used as biomarkers of aggressive PCa and prognosis, even serve as therapeutic targets. Furthermore, we should take into consideration that these genes are altered in progressing PCa rather than in primary PCa, which may robustly affect cancer progression.

Data availability

The data that support the findings of this study are openly available in Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo>), and the code for analysis in this article can be obtained by emailing the corresponding author.

Contributors

Guo-ping REN and Ji-li WANG conceived and designed the study. Ji-li WANG and Yan WANG collected the data and analyzed the data. Ji-li WANG wrote the original draft. Guo-ping REN and Yan WANG reviewed and edited the manuscript. All authors have read and approved the final manuscript. All authors have full access to all the data in the study and have responsibility for the integrity and security of the data.

Acknowledgments

The authors would like to thank The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) programs for providing high quality data.

Compliance with ethics guidelines

Ji-li WANG, Yan WANG, and Guo-ping REN declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Artavanis-Tsakonas S, Rand MD, Lake RJ, 1999. Notch signaling: cell fate control and signal integration in development. *Science*, 284(5415):770-776. <https://doi.org/10.1126/science.284.5415.770>
- Attard G, Parker C, Eeles RA, et al., 2016. Prostate cancer. *Lancet*, 387(10013):70-82. [https://doi.org/10.1016/S0140-6736\(14\)61947-4](https://doi.org/10.1016/S0140-6736(14)61947-4)
- Chang CC, Huang RL, Wang HC, et al., 2014. High methylation rate of *LMX1A*, *NKX6-1*, *PAX1*, *PTPRR*, *SOX1*, and *ZNF582* genes in cervical adenocarcinoma. *Int J Gynecol Cancer*, 24(2):201-209. <https://doi.org/10.1097/igc.000000000000054>

- Chen WQ, Zheng RS, Baade PD, et al., 2016. Cancer statistics in China, 2015. *CA Cancer J Clin*, 66(2):115-132. <https://doi.org/10.3322/caac.21338>
- D'Antonio JM, Ma CQ, Monzon FA, et al., 2008. Longitudinal analysis of androgen deprivation of prostate cancer cells identifies pathways to androgen independence. *Prostate*, 68(7):698-714. <https://doi.org/10.1002/pros.20677>
- Dennis G Jr, Sherman BT, Hosack DA, et al., 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, 4(5):P3. <https://doi.org/10.1186/gb-2003-4-5-p3>
- Duś-Szachniewicz K, Woźniak M, Nelke K, et al., 2015. Protein tyrosine phosphatase receptor R and Z1 expression as independent prognostic indicators in oral squamous cell carcinoma. *Head Neck*, 37(12):1816-1822. <https://doi.org/10.1002/hed.23835>
- Espinoza I, Pochampally R, Xing F, et al., 2013. Notch signaling: targeting cancer stem cells and epithelial-to-mesenchymal transition. *Oncotargets Ther*, 6:1249-1259. <https://doi.org/10.2147/ott.s36162>
- Gautier L, Cope L, Bolstad BM, et al., 2004. *affy*—analysis of *Affymetrix GeneChip* data at the probe level. *Bioinformatics*, 20(3):307-315. <https://doi.org/10.1093/bioinformatics/btg405>
- Gene Ontology Consortium, 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*, 32(S1):D258-D261. <https://doi.org/10.1093/nar/gkh036>
- Guo YZ, Sun HH, Wang XT, et al., 2018. Transcriptomic analysis reveals key lncRNAs associated with ribosomal biogenesis and epidermis differentiation in head and neck squamous cell carcinoma. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 19(9):674-688. <https://doi.org/10.1631/jzus.B1700319>
- Heidenreich A, Bastian PJ, Bellmunt J, et al., 2014. EAU guidelines on prostate cancer. Part II: treatment of advanced, relapsing, and castration-resistant prostate cancer. *Eur Urol*, 65(2):467-479. <https://doi.org/10.1016/j.eururo.2013.11.002>
- Kanehisa M, Goto S, 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*, 28(1):27-30. <https://doi.org/10.1093/nar/28.1.27>
- Kwon OJ, Zhang L, Wang JH, et al., 2016. Notch promotes tumor metastasis in a prostate-specific *Pten*-null mouse model. *J Clin Invest*, 126(7):2626-2641. <https://doi.org/10.1172/jci84637>
- Laczmanska I, Karpinski P, Bebenek M, et al., 2013. Protein tyrosine phosphatase receptor-like genes are frequently hypermethylated in sporadic colorectal cancer. *J Hum Genet*, 58(1):11-15. <https://doi.org/10.1038/jhg.2012.119>
- Li DM, Masiero M, Banham AH, et al., 2014. The Notch ligand *Jagged1* as a target for anti-tumor therapy. *Front Oncol*, 4:254. <https://doi.org/10.3389/fonc.2014.00254>
- Lin Y, Shen Z, Song X, et al., 2018. Comparative transcriptomic analysis reveals adriamycin-induced apoptosis via p53 signaling pathway in retinal pigment epithelial cells. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 19(12):895-909. <https://doi.org/10.1631/jzus.B1800408>
- Logothetis CJ, Lin SH, 2005. Osteoblasts in prostate cancer metastasis to bone. *Nat Rev Cancer*, 5(1):21-28. <https://doi.org/10.1038/nrc1528>
- Menigatti M, Cattaneo E, Sabates-Bellver J, et al., 2009. The protein tyrosine phosphatase receptor type R gene is an early and frequent target of silencing in human colorectal tumorigenesis. *Mol Cancer*, 8:124. <https://doi.org/10.1186/1476-4598-8-124>
- Munkley J, Lafferty NP, Kalna G, et al., 2015. Androgen-regulation of the protein tyrosine phosphatase PTPRR activates ERK1/2 signalling in prostate cancer cells. *BMC Cancer*, 15:9. <https://doi.org/10.1186/s12885-015-1012-8>
- Noordman YE, Jansen PAM, Hendriks WJAJ, 2006. Tyrosine-specific MAPK phosphatases and the control of ERK signaling in PC12 cells. *J Mol Signal*, 1:4. <https://doi.org/10.1186/1750-2187-1-4>
- Ritchie ME, Phipson B, Wu D, et al., 2015. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, 43(7):e47. <https://doi.org/10.1093/nar/gkv007>
- Santagata S, Demichelis F, Riva A, et al., 2004. JAGGED1 expression is associated with prostate cancer metastasis and recurrence. *Cancer Res*, 64(19):6854-6857. <https://doi.org/10.1158/0008-5472.can-04-2500>
- Schmitt I, Bitoun E, Manto M, 2009. PTPRR, cerebellum, and motor coordination. *Cerebellum*, 8(2):71-73. <https://doi.org/10.1007/s12311-009-0118-4>
- Sethi N, Dai XD, Winter CG, et al., 2011. Tumor-derived Jagged1 promotes osteolytic bone metastasis of breast cancer by engaging Notch signaling in bone cells. *Cancer Cell*, 19(2):192-205. <https://doi.org/10.1016/j.ccr.2010.12.022>
- Shannon P, Markiel A, Ozier O, et al., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498-2504. <https://doi.org/10.1101/gr.1239303>
- Shou JY, Ross S, Koeppen H, et al., 2001. Dynamics of Notch expression during murine prostate development and tumorigenesis. *Cancer Res*, 61(19):7291-7297.
- Siegel RL, Miller KD, Jemal A, 2019. Cancer statistics, 2019. *CA Cancer J Clin*, 69(1):7-34. <https://doi.org/10.3322/caac.21551>
- Su L, Song X, Xue Z, et al., 2018. Network analysis of microRNAs, transcription factors, and target genes involved in axon regeneration. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 19(4):293-304. <https://doi.org/10.1631/jzus.B1700179>
- Su PH, Lin YW, Huang RL, et al., 2013. Epigenetic silencing

- of PTPRR activates MAPK signaling, promotes metastasis and serves as a biomarker of invasive cervical cancer. *Oncogene*, 32(1):15-26.
<https://doi.org/10.1038/onc.2012.29>
- Su Q, Zhang B, Zhang L, et al., 2017. Jagged1 upregulation in prostate epithelial cells promotes formation of reactive stroma in the Pten null mouse model for prostate cancer. *Oncogene*, 36(5):618-627.
<https://doi.org/10.1038/onc.2016.232>
- Sun YT, Wang BE, Leong KG, et al., 2012. Androgen deprivation causes epithelial-mesenchymal transition in the prostate: implications for androgen-deprivation therapy. *Cancer Res*, 72(2):527-536.
<https://doi.org/10.1158/0008-5472.can-11-3004>
- Szklarczyk D, Morris JH, Cook H, et al., 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*, 45(D1):D362-D368.
<https://doi.org/10.1093/nar/gkw937>
- Tang ZF, Li CW, Kang BX, et al., 2017. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*, 45(W1):W98-W102.
<https://doi.org/10.1093/nar/gkx247>
- Terada N, Shimizu Y, Kamba T, et al., 2010. Identification of EP4 as a potential target for the treatment of castration-resistant prostate cancer using a novel xenograft model. *Cancer Res*, 70(4):1606-1615.
<https://doi.org/10.1158/0008-5472.can-09-2984>
- Watson PA, Arora VK, Sawyers CL, 2015. Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nat Rev Cancer*, 15(12):701-711.
<https://doi.org/10.1038/nrc4016>
- Weinstein JN, Collisson EA, Mills GB, et al., 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*, 45(10):1113-1120.
<https://doi.org/10.1038/ng.2764>
- Woźniak M, Gamian E, Łączmańska I, et al., 2014. Immunohistochemical and Western blot analysis of two protein tyrosine phosphatase receptors, R and Z1, in colorectal carcinoma, colon adenoma and normal colon tissues. *Histol Histopathol*, 29(5):635-639.
<https://doi.org/10.14670/HH-29.10.635>
- Yong T, Sun A, Henry MD, et al., 2011. Down regulation of CSL activity inhibits cell proliferation in prostate and breast cancer cells. *J Cell Biochem*, 112(9):2340-2351.
<https://doi.org/10.1002/jcb.23157>
- Zhu H, Zhou XC, Redfield S, et al., 2013. Elevated Jagged-1 and Notch-1 expression in high grade and metastatic prostate cancers. *Am J Transl Res*, 5(3):368-378.
- Zhu H, Li Y, Wang M, et al., 2019. Analysis of cardiovascular disease-related NF- κ B-regulated genes and microRNAs in TNF α -treated primary mouse vascular endothelial cells. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 20(10):803-815.
<https://doi.org/10.1631/jzus.B1800631>

List of electronic supplementary materials

Table S1 Clinical characteristics of samples in this study
 Fig. S1 Expression of the *PTPRR* in Grasso PCa

中文概要

题目: 通过生物信息学方法鉴定 *PTPRR* 和 *JAG1* 为去势抵抗性前列腺癌 (CRPC) 的关键基因

目的: 鉴定去势抵抗性前列腺癌 (CRPC) 的关键基因。

创新点: (1) 结合多个数据库数据, 运用生物信息学方法鉴定 CRPC 的关键基因; (2) 首次报道 *PTPRR* 可能在 CRPC 里起关键作用。

方法: 通过下载三个 GEO 数据库的 mRNA 微阵列数据, 分析 CRPC 和激素敏感前列腺癌之间的基因差异, 对筛选出的差异基因进行功能富集分析和蛋白质间相互作用分析, 最终筛选出两个有重要功能的差异基因 (*PTPRR* 和 *JAG1*)。通过在多个其他数据库中进行表达量验证和生存分析, 进一步证明这些基因的重要作用。

结论: *PTPRR* 和 *JAG1* 在 CRPC 中显著增高, 并与预后差相关。因此, 这两个基因有可能作为 CRPC 的诊断和预后的生物标志物。

关键词: 生物信息学; *PTPRR*; *JAG1*; 差异表达基因; 激素抵抗前列腺癌; 功能富集