



Identification of Metabolism-Related Gene-Based Subgroup in Prostate Cancer

Guopeng Yu^{1†}, Bo Liang^{2†}, Keneng Yin^{3†}, Ming Zhan¹, Xin Gu¹, Jiangyi Wang¹, Shangqing Song¹, Yushan Liu^{1*}, Qing Yang^{1*}, Tianhai Ji^{3,4*} and Bin Xu^{1*}

OPEN ACCESS

Edited by:

Rong Na,
The University of Hong Kong, Hong Kong SAR, China

Reviewed by:

Da Huang,
Shanghai Jiao Tong University School of Medicine, China
Fubo Wang,
Second Military Medical University, China

*Correspondence:

Bin Xu
chxb2004@126.com
Tianhai Ji
skysea_ji@sina.com
Qing Yang
dryangq@gmail.com
Yushan Liu
yushandoctor@163.com

[†]These authors have contributed equally to this work

Specialty section:

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

Received: 31 March 2022

Accepted: 19 May 2022

Published: 16 June 2022

Citation:

Yu G, Liang B, Yin K, Zhan M, Gu X, Wang J, Song S, Liu Y, Yang Q, Ji T and Xu B (2022) Identification of Metabolism-Related Gene-Based Subgroup in Prostate Cancer. *Front. Oncol.* 12:909066. doi: 10.3389/fonc.2022.909066

¹ Department of Urology, Shanghai Ninth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China, ² The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China, ³ 174 Clinical College, Anhui Medical University, Hefei, China, ⁴ Department of Pathology, Shanghai Ninth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

Prostate cancer is still the main male health problem in the world. The role of metabolism in the occurrence and development of prostate cancer is becoming more and more obvious, but it is not clear. Here we firstly identified a metabolism-related gene-based subgroup in prostate cancer. We used metabolism-related genes to divide prostate cancer patients from The Cancer Genome Atlas into different clinical benefit populations, which was verified in the International Cancer Genome Consortium. After that, we analyzed the metabolic and immunological mechanisms of clinical beneficiaries from the aspects of functional analysis of differentially expressed genes, gene set variation analysis, tumor purity, tumor microenvironment, copy number variations, single-nucleotide polymorphism, and tumor-specific neoantigens. We identified 56 significant genes for non-negative matrix factorization after survival-related univariate regression analysis and identified three subgroups. Patients in subgroup 2 had better overall survival, disease-free interval, progression-free interval, and disease-specific survival. Functional analysis indicated that differentially expressed genes in subgroup 2 were enriched in the xenobiotic metabolic process and regulation of cell development. Moreover, the metabolism and tumor purity of subgroup 2 were higher than those of subgroup 1 and subgroup 3, whereas the composition of immune cells of subgroup 2 was lower than that of subgroup 1 and subgroup 3. The expression of major immune genes, such as CCL2, CD274, CD276, CD4, CTLA4, CXCR4, IL1A, IL6, LAG3, TGFB1, TNFRSF4, TNFRSF9, and PDCD1LG2, in subgroup 2 was almost significantly lower than that in subgroup 1 and subgroup 3, which is consistent with the results of tumor purity analysis. Finally, we identified that subgroup 2 had lower copy number variations, single-nucleotide polymorphism, and neoantigen mutation. Our systematic study established a metabolism-related gene-based subgroup to predict outcomes of prostate cancer patients, which may contribute to individual prevention and treatment.

Keywords: prostate cancer, metabolism, immune, non-negative matrix factorization (NMF), The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC)

INTRODUCTION

Prostate cancer (PCa) is the most common urological cancer among men in the United States (1). It is reported that in 2021, there were an estimated 250,000 new cases and 34,000 deaths (2). The 5-year prevalence of PCa is the highest globally, and its age-standardized mortality is the sixth highest (1). The clinical diagnosis of PCa mainly depends on digital rectal examination, serum prostate-specific antigen, and imaging examination (3). Usually, prostate biopsy for pathological examination is also required. Gleason score can help to evaluate the malignancy of PCa. Although radical prostatectomy has become the main strategy for resection of localized primary prostate tumors, more than one in five PCa patients inevitably progress to the advanced stage with a poor prognosis within 10 years (4, 5). Due to the high rate of bone metastasis, PCa patients shared an unfavorable prognosis (6). Other treatment methods, such as hormone therapy and chemoradiotherapy, are not satisfactory for the prognosis (7). In addition, PCa has high heterogeneity, which results in different prognoses of patients after treatment (8). Therefore, there is a need to deeply understand the potential mechanism leading to PCa progression and metastasis, and select specific subtypes to find the population who benefit the most.

The cell proliferation state in tumor progression involves the corresponding changes of cell metabolism (9), and metabolic reprogramming is considered to be a basic feature of cancer cells (9). Metabolism in tumors, such as the Warburg effect and glutamine metabolism, is significantly different from that in normal tissues (10, 11). There is increasing evidence that metabolic abnormalities are associated with poor prognosis of many tumor types (12). Fortunately, the screening of metabolic biomarkers can specifically detect abnormal changes in organisms to prevent malignant diseases with pathophysiological characteristics (13). The metabolic markers have been well displayed in a variety of tumors, such as hepatocellular carcinoma (14), colorectal cancer (15), endometrial cancer (16), and clear cell renal cell carcinoma (17), but research in PCa is still relatively scarce. Therefore, it is of great clinical significance to find a new metabolic marker to predict the prognosis of PCa.

In the present study, we systemically analyzed the profile of metabolism-related genes from The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC). All patients could be grouped into three subgroups with different prognoses through non-negative matrix factorization (NMF) based on TCGA, which was validated in ICGC. Then we obtained differentially expressed genes (DEGs) and chronologically conducted functional analysis, gene set variation analysis (GSVA), and immune-related comprehensive analysis (Figure S1).

MATERIALS AND METHODS

Acquisition and Processing of Raw Data

We downloaded the RNA-sequence data, clinical information, survival data, and somatic mutation data of TCGA prostate adenocarcinoma (PRAD) from UCSC Xena ([\[xenabrowser.net/datapages/\]\(https://xenabrowser.net/datapages/\)\). According to clinical information corresponding to the sample, we only select these samples whose primary diagnosis is “adenocarcinoma, NOS” and sample type is “primary tumor” follow-up analysis. After the expression profile was reannotated, the expression data of all mRNAs were selected according to the human genome information contained in the HUGO Gene Nomenclature Committee \(HGNC\) database \(18\), and a new expression profile was obtained.](https://</p></div><div data-bbox=)

Then, we downloaded the gene expression files, sample information, and clinical information corresponding to PRAD-CA from the ICGC database, removed the normal samples (only PRAD samples were included), and sorted out the new expression profile data according to the HGNC database for subsequent verification and analysis.

Screening of Metabolism-Related PRAD Genes

We cross-referenced maps of metabolic pathways with the Kyoto Encyclopedia of Genes and Genomes database to compile a comprehensive list of 2,752 genes encoding all known human metabolic enzymes and transporters (19). After intersecting with TCGA PRAD mRNA expression profile, we obtained the metabolism-related genes in PRAD. After that, we removed the genes with median absolute deviation (MAD) ≤ 0.5 to obtain metabolism-related PRAD genes.

NMF and Survival Analysis

We performed survival-related univariate regression analysis on metabolism-related PRAD genes and selected the significant genes ($P < 0.05$) for subsequent NMF analysis. Based on the expression profile of these significant genes, we performed NMF analysis through the NMF package (20). Taking the front point of the maximum change of the cophenetic correlation with K as the best rank for NMF analysis, TCGA PRAD metabolism-related subgroups were obtained, and then dimensionality reduction visualization was carried out by using principal component analysis (PCA) and the t -distributed stochastic neighbor embedding (t -SNE) unsupervised clustering method was used to view the characteristics between different metabolic subgroups.

Moreover, we applied the expression profile of the same significant genes in ICGC to conduct NMF analysis using the same best rank as mentioned above. Then we used the SubMap module (21) of GenePattern (22) to map the subgroups obtained from TCGA and ICGC.

After obtaining the results of metabolism-related subgroups in TCGA and ICGC, we used the *survival* package (<https://CRAN.R-project.org/package=survival>) to analyze the survival of different metabolic subgroups in the two datasets, as described previously (23), and compared the prognosis of the corresponding subgroups in the two different datasets.

Identification of DEGs

In order to obtain the possible unique molecular biological functions of different metabolic subgroups, we analyzed the differences among three subgroups obtained from TCGA data. Differential analysis was performed based on the *limma*

package (24), and the screening threshold of DEGs was $|\log_2^{\text{Fold Change}}| > 1$ and adjusted $P < 0.01$.

Then, we used *jveen*, a flexible tool, to cross analyze the results of the three groups of DEGs (25) and obtain the unique DEGs of each subgroup relative to other subgroups, and then visualized the corresponding expression of these DEGs with a cluster heatmap.

Functional Analysis

We used the Metascape database to enrich and analyze the unique DEGs of each subgroup (26) to explore the possible molecular biological functions, as described previously (27).

GSVA

GSVA is a non-parametric and unsupervised gene set enrichment method that can estimate the score of a certain pathway or signature based on transcriptomic data (28). Similarly, we employed 2,752 metabolism-related genes from the previous study (19) to conduct GSVA using the *gsva* package (28). Then we got the score of each sample under 113 metabolic items and then conducted ANOVA among multiple subtypes based on the metabolic score. $P < 0.01$ was considered significant. For the significant metabolic items, we further conducted a Tukey posttest to judge the differences among different subgroups under the corresponding items. For each metabolic item, we selected the metabolic items with $|\text{diff}| > 0.2$ and adjusted $P < 0.01$ as the items with significant differences among different subgroups.

Tumor Purity Analysis

In the tumor microenvironment, immune cells and stromal cells are two main types of tumor cells. ESTIMATE uses the expression profiles to predict the score of stromal cells and immune cells and then predicts the content of these two cells (29). Therefore, the tumor purity in each tumor sample can be calculated; that is, if there are many stromal cells and immune cells, the tumor purity is low, and on the contrary, the tumor purity is high. Here, we analyzed the tumor purity of TCGA based on the *estimate* package (<https://R-Forge.R-project.org/projects/estimate>), as described previously (30).

Tumor Immune Microenvironment Analysis

CIBERSORT is a method to deconvolute the expression matrix of human immune cell subtypes based on the principle of linear support vector regression (31). It is mostly used for gene expression profiles, and the deconvolution analysis of unknown mixture and expression profiles containing similar cell types is better than other methods. Based on CIBERSORT, we obtained the composition of 22 immune cells in TCGA and statistically analyzed the immune microenvironment of different subtypes.

Mutational Cancer Driver and Immune Gene Analysis

The Integrative OncoGenomics (IntOGen) pipeline is an implementation to obtain the compendium of mutational cancer drivers. Its application to somatic mutations of more than 28,000 tumors of 66 cancer types reveals 568 cancer genes and points toward their mechanism of tumorigenesis (32). We

downloaded the cancer driver genes of PRAD from IntOGen and analyzed the main tumor driver gene (CGC_CANCER_Gene = TRUE) among subgroups. At the same time, we also analyzed the expression of major immune genes among subgroups.

CNV, SNP, and Tumor-Specific Neoantigen Analysis

We downloaded the CNV data of TCGA PRAD from the GDAC Firehose database (<http://gdac.broadinstitute.org/>), as described previously (33), and then used the Gistic (version 2.0) module of GenePattern to analyze the CNV of each subgroup. The *maftools* package was used for visualization (34).

At the same time, we downloaded the PRAD somatic SNP results under the varscan processing flow from TCGA. After calculating the TMB, we used the *maftools* package (34) to statistically visualize the top 20 mutant genes and mutational cancer driver genes.

TSNadb is a comprehensive tumor-specific neoantigen database based on pan-cancer immunogenomic analysis of somatic mutation data and human leukocyte antigen (HLA) allele information for 16 tumor types with 7,748 tumor samples from TCGA and The Cancer Immunome Atlas (TCIA) (35). We downloaded the tumor-specific neoantigens in PRAD and then analyzed the mutations carried by the tumor-specific neoantigens in the three subgroups.

RESULTS

Patient Characteristics

Through acquisition and processing of raw data from TCGA and ICGC, we obtained 485 and 144 samples, respectively. The patient characteristics are listed in **Table 1**. After HGNC validation, we obtained 18,400 and 18,570 mRNAs in TCGA and ICGC, respectively (**Table S1**).

Metabolism-Related PRAD Genes

We crossed 2,752 metabolism-related genes with 18,400 genes in TCGA PRAD to obtain 2,579 metabolism-related genes in PRAD. After further screening, 2,243 metabolism-related PRAD genes were left.

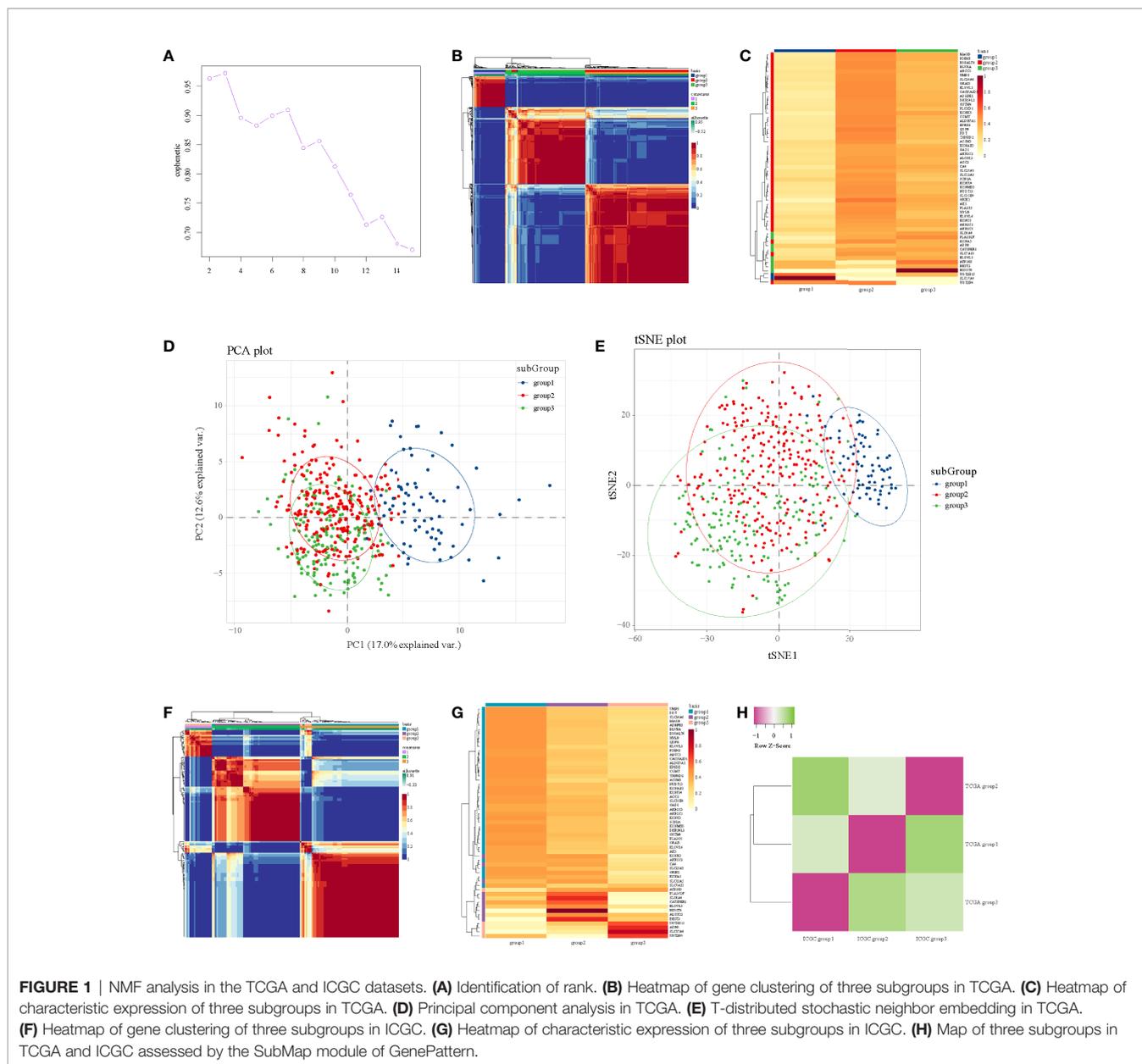
NMF Analysis

We identified 56 significant genes for NMF after survival-related univariate regression analysis (**Table S2**). From **Figure 1A**, we identified three as the best rank for NMF analysis. Then, we divided 56 significant genes into three subgroups. A heatmap of the gene clustering method in TCGA showed significant differences in expression levels (**Figure 1B**). We also generated a heatmap to show the characteristic expression of 56 significant genes in the three subgroups (**Figure 1C**). The clinical relevance of the three subgroups is shown in **Figure S2**. The principal component analysis unsupervised clustering method indicated that the samples of different subgroups had significant characteristics, while subgroups 2 and subgroups 3 had certain

TABLE 1 | Characteristics of patients in TCGA and ICGC datasets.

Variable	TCGA (N = 485)	ICGC (N = 144)
Age (years)		
≤60	219 (45%)	59 (41%)
>60	266 (55%)	85 (59%)
Race		
Asian	12 (2%)	Not applicable
Black or African American	56 (12%)	Not applicable
White	404 (83%)	Not applicable
Others	13 (3%)	Not applicable
Gleason score		
6	46 (9%)	29 (20%)
7	245 (51%)	80 (56%)
8	59 (12%)	9 (6%)
9	131 (27%)	2 (1%)
10	4 (1%)	Not applicable
Other	0 (0%)	24 (17%)
Prostate-specific antigen (ng/ml)		
<4	402 (83%)	Not applicable
≥4	26 (5%)	Not applicable
No detection	57 (12%)	Not applicable
T stage		
T1	0 (0%)	83 (58%)
T2	188 (39%)	61 (42%)
T3	280 (58%)	0 (0%)
T4	10 (2%)	0 (0%)
No detection	7 (1%)	0 (0%)
N stage		
N0	442 (91%)	Not applicable
N1	3 (1%)	Not applicable
No detection	40 (8%)	Not applicable
M stage		
M0	336 (69%)	Not applicable
M1	77 (16%)	Not applicable
No detection	72 (15%)	Not applicable
Tumor status		
Tumor free	338 (70%)	Not applicable
With tumor	88 (18%)	Not applicable
No detection	59 (12%)	Not applicable
New event		
No	357 (74%)	Not applicable
Yes	104 (21%)	Not applicable
No detection	24 (5%)	Not applicable
Radiation therapy		
No	383 (79%)	Not applicable
Yes	59 (12%)	Not applicable
No detection	43 (9%)	Not applicable
Primary therapy outcome		
Complete response	331 (68%)	Not applicable
Progressive disease	26 (5%)	Not applicable
Partial response	40 (8%)	Not applicable
Stable disease	29 (6%)	Not applicable
Other	59 (12%)	Not applicable
Residual tumor		
R0	206 (42%)	Not applicable
R1	144 (30%)	Not applicable
R2	5 (1%)	Not applicable
No detection	130 (27%)	Not applicable
Zone of origin		
Central zone	4 (1%)	Not applicable
Overlapping/multiple zones	124 (26%)	Not applicable
Peripheral zone	134 (28%)	Not applicable
Transition zone	7 (1%)	Not applicable
No detection	216 (45%)	Not applicable

Data were shown as n (%).



similarities (Figure 1D). The same results were obtained by t-distributed stochastic neighbor embedding (t-SNE) (Figure 1E).

Then, we used the data from ICGC to verify these three subgroups in TCGA; we also used three as the rank for NMF analysis of the expression profile of 56 significant genes in ICGC to obtain three subgroups, and the internal consistency of the three subgroups was also good (Figure 1F). Figure 1G shows the characteristic expression of 56 significant genes in the three subgroups in ICGC. Through mapping the three subgroups from TCGA and ICGC, we confirmed their corresponding relationships (Figure 1H). Detailly, TCGA subgroup 3 corresponded to ICGC subgroup 2, TCGA subgroup 1 corresponded to ICGC subgroup 3, and TCGA subgroup 2 corresponded to ICGC subgroup 1 (Figure 1H). Importantly, the corresponding relationships were

consistent with the characteristic expression of 56 significant genes in the subgroups (Figures 1C, G), indicating that our submap matching analysis results of subgroups in the two datasets were good.

Survival Analysis

Then, we performed survival analysis under subgroups in both TCGA and ICGC to explore the prognosis. According to TCGA data, the three subgroups had statistically significant overall survival differences ($P = 0.0017$, Figure 2A), and we can observe that the overall survival of subgroup 1 was the worst, subgroup 2 was the second, subgroup 3 has a better overall survival, and the overall survival of subgroup 2 and subgroup 3 was the same. Moreover, patients in subgroup 1 shared an unfavorable disease-free interval

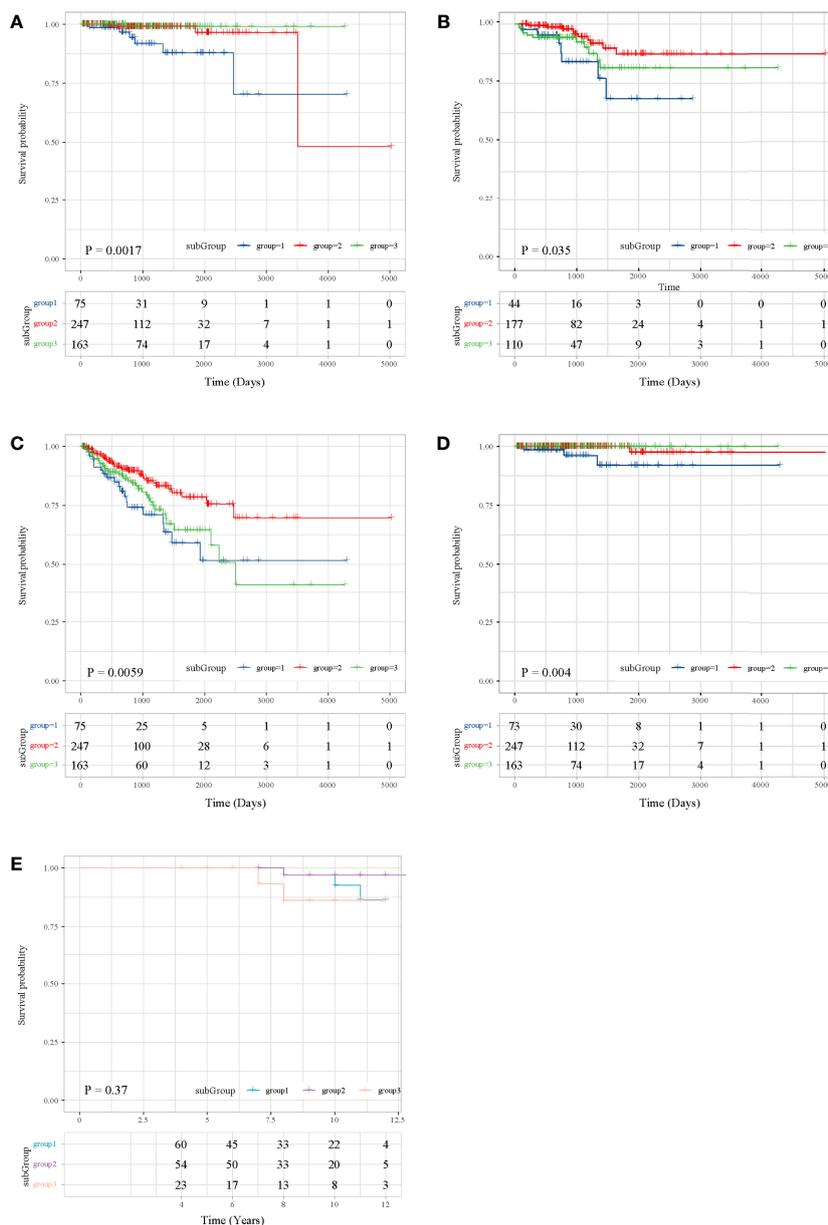


FIGURE 2 | Survival analysis in the TCGA and ICGC datasets. **(A)** OS in TCGA. **(B)** DFI in TCGA. **(C)** PFI in TCGA. **(D)** DFS in TCGA. **(E)** OS in ICGC.

($P = 0.035$, **Figure 2B**), progression-free interval ($P = 0.0059$, **Figure 2C**), and disease-specific survival ($P = 0.004$, **Figure 2D**). In ICGC, the overall survival of subgroup 3 was the worst, subgroup 1 was the second, subgroup 2 was the best, and the overall survival of subgroup 1 and subgroup 2 was the same (**Figure 2E**). These results were consistent with the results of subgroup matching of the two datasets (**Figure 1H**).

Functional Analysis of DEGs

Through multi-subgroup DEG screening, we obtained 88 DEGs in subgroup 1, 15 DEGs in subgroup 2, and 166 DEGs in subgroup 3 (**Figure 3A**). Moreover, we found that most of the DEGs in

subgroup 1 and subgroup 3 were significantly up-regulated, whereas most of the DEGs in subgroup 2 were down-regulated (**Figure 3B**). Later functional analysis indicated that these DEGs in the three subgroups were enriched with some metabolism-related biological processes and pathways (**Figures 3C–E**).

GSVA

Through preliminary screening, we obtained 15 significant metabolic items (**Figure 4A**). Among them, for most metabolic pathways, the metabolism of subgroup 2 was higher than that of subgroup 1 and subgroup 3. Detailly, cyclooxygenase arachidonic acid metabolism and ascorbate and aldrate metabolism were the



FIGURE 3 | Functional analysis of DEGs in three subgroups. **(A)** Venn diagram of DEGs in three subgroups. **(B)** Heatmap of DEGs in three subgroups. **(C)** Functional analysis of DEGs in subgroup 1. **(D)** Functional analysis of DEGs in subgroup 2. **(E)** Functional analysis of DEGs in subgroup 3.

highest in subgroup 1, whereas prostanoid biosynthesis; propanote metabolism; valine, leucine, and isoleucine degradation; beta-alanine metabolism; glycosphingolipid biosynthesis; other glycan degradation; and caffeine metabolism were the lowest in subgroup 1 (Figure 4B).

Tumor Purity and Immune Microenvironment Analysis

The results indicated that the tumor purity of subgroup 2 was significantly higher than that of subgroup 1 and subgroup 3, and the composition of immune cells of subgroup 2 was lower than that of subgroup 1 and subgroup 3 (Figures 5A, B).

Through CIBERSORT, we found that the main immune cell type is T cells CD4 memory resetting, followed by macrophage M2, plasma cells, and B cells naïve in TCGA (Figure 5C). T cells CD4 memory resting, which were the main components of PRAD, were significantly lower than subgroup 1 and subgroup 3 (Figure 5D), which was consistent with tumor purity. However, in subgroup 2, there were significantly more plasma cells than in subgroup 1 and subgroup 3 (Figure 5D).

Mutational Cancer Driver and Immune Gene Analysis

From IntOGen, we obtained 11 mutational cancer driver genes in PRAD. Among these genes, CDKN2A was down-regulated in

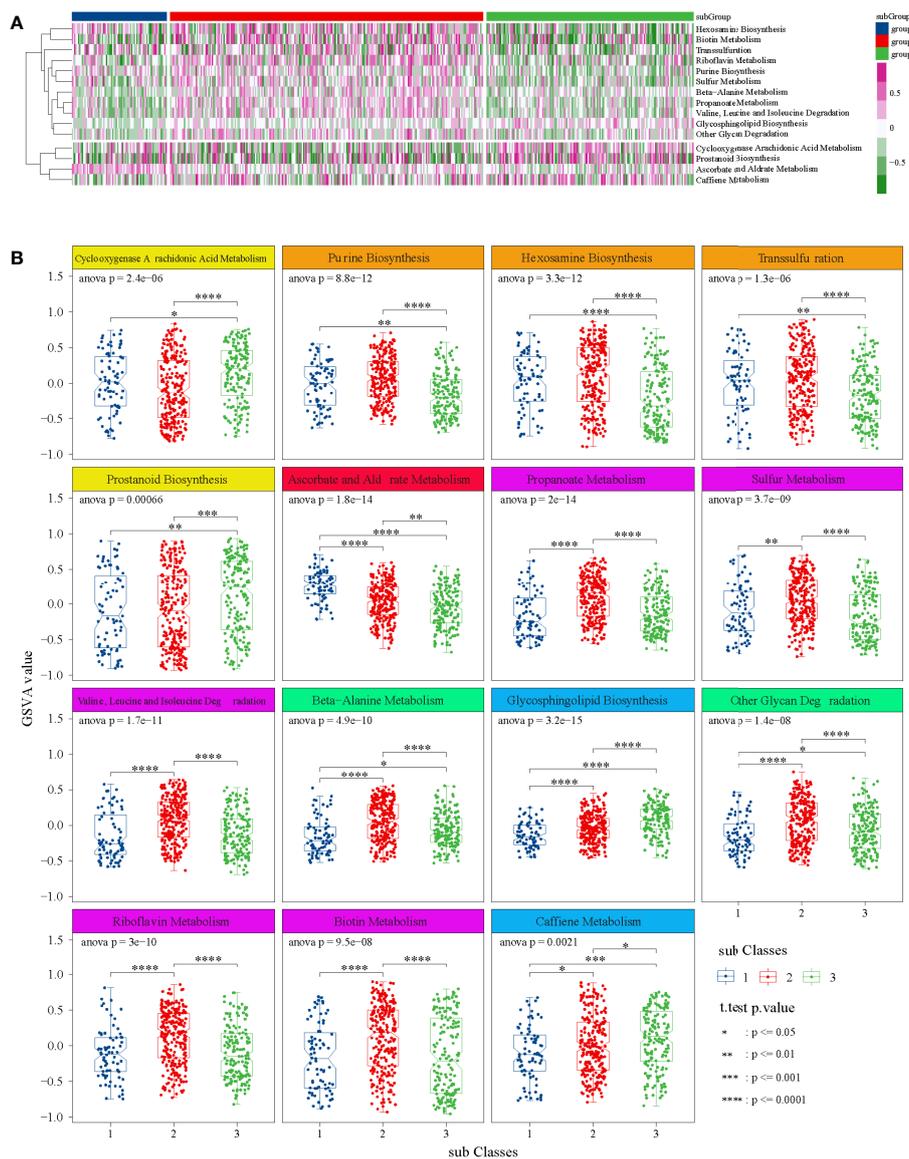


FIGURE 4 | GSEA in TCGA dataset. **(A)** Heatmap of 15 significant metabolic items in three subgroups. **(B)** Box diagrams of 15 significant metabolic items in three subgroups. ANOVA test was performed among groups, and t-test was performed between the two groups.

subgroup 1, BRCA2 was down-regulated in subgroup 2, and BNF43 was up-regulated in subgroup 2 (Figure S3A). Moreover, we demonstrated that the expression of major immune genes, such as CCL2, CD274, CD276, CD4, CTLA4, CXCR4, IL1A, IL6, LAG3, TGFB1, TNFRSF4, TNFRSF9, and PDCD1LG2, in subgroup 2 was almost significantly lower than that in subgroup 1 and subgroup 3 (Figure S3B), which is consistent with the results of tumor purity analysis.

CNV, SNP, and Tumor-Specific Neoantigen Analysis

Each area of CNV is assigned a G-score that considers the amplitude of the alteration as well as the frequency of its

occurrence across samples (36). Subgroup 1 had amplifications of 8q24.21, subgroup 2 had amplifications of 13q33.3, 3q26.2, and 8q22.1, and subgroup 3 had amplifications of 13q12.11, 8p12, and 8q22.3 (Figure 6A and Table S3). Interestingly, all subgroups had deletions of 10q23.31, 3p13, 5q11.2, 5q21.1, 6q14.3, 13q14.13, 17p13.1, and 21q22.3 (Figure 6A and Table S3). In addition, subgroup 1 had the highest copy number amplification (Figure 6B) and copy number deletion (Figure 6C). Moreover, the tumor-specific neoantigens of subgroup 1 carried more mutations than subgroup 2 and subgroup 3 (Figure 6D).

Through SNP analysis, missense mutation was the most frequent variant classification and C>T was the most frequent single-nucleotide variant (Figure 7A). From the oncoplots of top

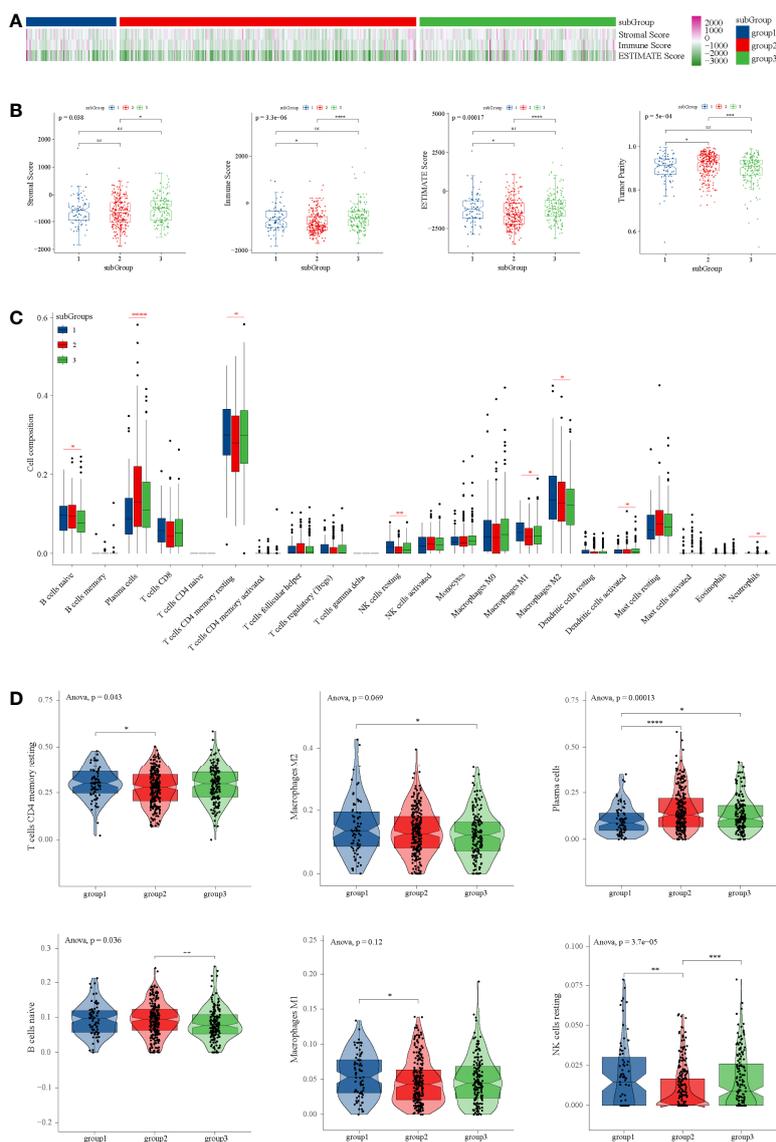


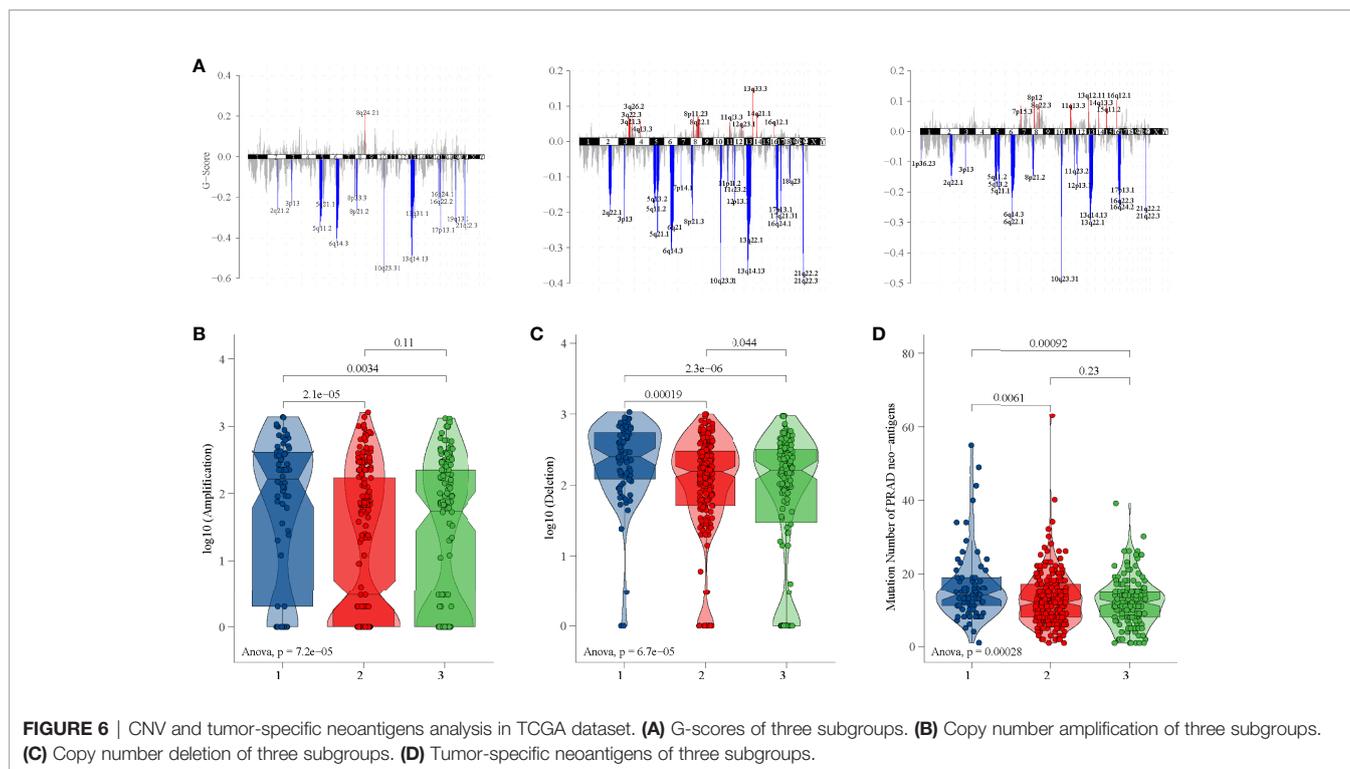
FIGURE 5 | Tumor purity and immune microenvironment analysis in TCGA dataset. **(A)** Heatmap of tumor purity analysis in three subgroups. **(B)** Box diagrams tumor purity analysis in three subgroups. **(C)** Composition of 22 immune cells in TCGA. **(D)** Immune cell types with significant differences among subgroups. ANOVA test was performed among groups, and t-test was performed between the two groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ and ns $P \geq 0.05$.

20 mutant genes (**Figure 7B**) and mutational cancer driver genes (**Figure 7C**), we found that TP53, FAT3, LRP1B, ATM, KMT2C, and KMT2D from mutant genes were also recognized as cancer driver genes.

DISCUSSION

In recent years, the heterogeneity of PCa has been an important topic of research (37). Exploring new tumor subtypes, especially combined metabolism, is an effective way to study their heterogeneity and thus provides insights for

clinicians to conduct more accurate clinical evaluation. Bioinformatics analysis based on database has increasingly shown its superiority and clinical applicability (23, 38). Here, we used metabolism-related genes to conduct consensus clustering among large-scale PCa patients and finally identified three PCa subgroups. These subgroups were validated by an external clinical patient cohort from ICGC. We conducted GSEA to enrich pathways. ESTIMATE and CIBERSORT algorithms were used to conduct an integrative analysis of immune scores and immune cells in the different subgroups. Finally, cancer driver genes, CNV, SNP, and tumor-specific neoantigens were analyzed. Our results



from this study might reveal the molecular mechanism of metabolism-induced PCa.

Through survival analysis, we found that the prognosis of patients in subgroup 1 was significantly worse than that of the other two subgroups in TCGA, and there was no significant difference in prognosis among other two subgroups. Moreover, the prognosis of patients in subgroup 3 (corresponded to TCGA subgroup 1) was significantly worse than that of the other two subgroups in ICGC and there was no significant difference in prognosis among other two subgroups. Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) showed a clear boundary between these subgroups in both TCGA and ICGC.

Steroidogenic enzymes are essential for PCa development (39). 17β HSD2 decreased potent androgen production by converting testosterone or dihydrotestosterone to each of their upstream precursors (40), which might provide new strategies for clinical management. On the other hand, chemokine signaling regulates tumor metastasis (41). CXCL12, a member of the chemokine family, and its receptor, CXCR4, are key mediators of PCa bone metastasis (42). What is more, the tyrosine 190 and 211 phosphorylation of proliferation cell nuclear antigen is a frequent event in advanced prostate cancer (43, 44). Strikingly, we found that steroid hormone biosynthesis, regulation of chemokine production, and peptidyl-tyrosine phosphorylation were enriched in subgroup 1 in TCGA. Cyclooxygenase is a rate-limiting enzyme involved in the cyclooxygenase metabolic pathway of arachidonic acid, which can catalyze the conversion of arachidonic acid to prostaglandins (45). A systematic review indicated that an 8.75-month increase

in progression-free survival and an improved trend in overall survival in the cancers received ascorbate (46). Interestingly, our study identified that cyclooxygenase arachidonic acid metabolism and ascorbate and aldurate metabolism were the highest in subgroup 1. Altogether, our results explain the poor prognosis of this particular population.

Patients in subgroup 1 had the highest tumor purity, causing worse prognosis. We also found that T cells CD4 memory resetting, macrophages M2, macrophages M1, B cells naïve, and NK cells resetting infiltrated significantly more in subgroup 1. Previous data showed that PRAD patients with high numbers of M2 macrophages in the tumor environment had increased odds of dying (47). These cells could promote PCa progression by promoting immunosuppressive responses (48, 49). Through CNV analysis, we found that 8q24.21, strongly associated with risk of tumors (50, 51), was amplified in subgroup 1. Some long non-coding RNAs, such as CCAT1 (52, 53), PVT1 (54), PRNCR1 (55), and PCAT1 (56), in 8q24.21 have an influence in oncogenesis of PCa. The tumor-specific neoantigens might contribute to the immunogenic phenotype (57). In prostate cancer, inactivating CDK12 mutations produces tumor-specific neoantigens and possibly sensitivity to immunotherapy (58). Vaccines that target tumor-specific neoantigens have the potential to induce robust antitumor responses (59). Sipuleucel-T, a neoantigen vaccine, prolonged overall survival among men with metastatic castration-resistant prostate cancer (60). In the present study, we confirmed more mutations of tumor-specific neoantigens in subgroup 1, which may be one of the reasons for the poor prognosis.

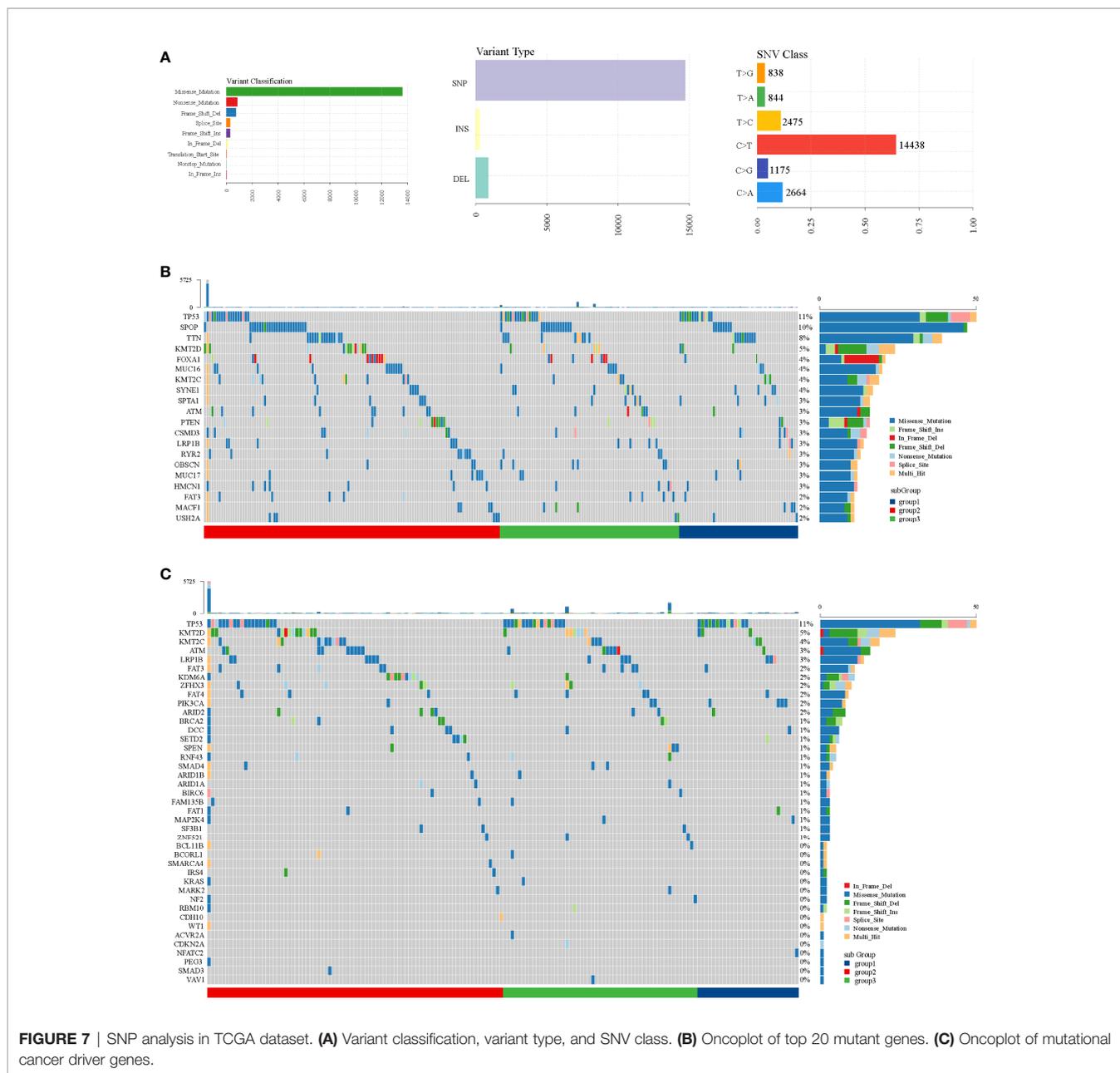


FIGURE 7 | SNP analysis in TCGA dataset. **(A)** Variant classification, variant type, and SNV class. **(B)** Oncoplot of top 20 mutant genes. **(C)** Oncoplot of mutational cancer driver genes.

There are some limitations in this study. First, the available public datasets in our study were from TCGA and ICGC, two independent data platforms. Although the results can be well verified, there is inevitable selection bias. We will further validate our results in the inpatients at our hospitals. Second, limited by the survival data in ICGC, we could not conduct multiple survival analyses like TCGA; we can only do OS analysis. Moreover, although we used different datasets as the training cohort and validation cohort respectively, the sample size of each cohort is still small, and large-sample data are needed to verify our findings. Finally, more solid experimental studies are

needed to verify DEGs and their immune and molecular biological mechanism.

CONCLUSION

Collectively, the metabolic mechanism of PCa was systematically explored, providing associations with mutational burden and immune infiltrations. PCa-related signatures prove to have advantages in predicting prognosis and can be used as a good molecular classifier to find different metabolic types. The relevant

findings will need further basic experiments and even clinical trials to be corroborated in the future before they can be further applied in clinical practice.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

GY, BL, and KY contributed equally as first authors. BX, TJ, QY, and YL contributed equally as senior authors. GY, BX, and MZ conceived, designed, or planned the idea. XG, JW, and SS contributed to the language editing of the manuscript. GY, BL, and KY analyzed the data. All authors collected the data. All authors interpreted the results. GY and BL wrote the manuscript. BX, TJ, QY, and YL revised the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

- Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590
- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654
- Pernar CH, Ebot EM, Wilson KM, Mucci LA. The Epidemiology of Prostate Cancer. *Cold Spring Harb Perspect Med* (2018) 8(12):a030361. doi: 10.1101/cshperspect.a030361
- Litwin MS, Tan H-J. The Diagnosis and Treatment of Prostate Cancer: A Review. *JAMA* (2017) 317(24):2532–42. doi: 10.1001/jama.2017.7248
- Bill-Axelsson A, Holmberg L, Garmo H, Taari K, Busch C, Nordling S, et al. Radical Prostatectomy or Watchful Waiting in Prostate Cancer — 29-Year Follow-Up. *N Engl J Med* (2018) 379(24):2319–29. doi: 10.1056/NEJMoa1807801
- Rawla P. Epidemiology of Prostate Cancer. *World J Oncol* (2019) 10(2):63–89. doi: 10.14740/wjon1191
- Rodríguez-Ruiz ME, Perez-Gracia JL, Rodríguez I, Alfaro C, Oñate C, Pérez G, et al. Combined Immunotherapy Encompassing Intratumoral Poly-Icld, Dendritic-Cell Vaccination and Radiotherapy in Advanced Cancer Patients. *Ann Oncol Off J Eur Soc Med Oncol* (2018) 29(5):1312–9. doi: 10.1093/annonc/mdy089
- Segura-Moreno YY, Sanabria-Salas MC, Varela R, Mesa JA, Serrano ML. Decoding the Heterogeneous Landscape in the Development Prostate Cancer. *Oncol Lett* (2021) 21(5):376. doi: 10.3892/ol.2021.12637
- Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
- Hsu PP, Sabatini DM. Cancer Cell Metabolism: Warburg and Beyond. *Cell* (2008) 134(5):703–7. doi: 10.1016/j.cell.2008.08.021
- Qu W, Oya S, Lieberman BP, Ploessl K, Wang L, Wise DR, et al. Preparation and Characterization of L-[5-11c]-Glutamine for Metabolic Imaging of Tumors. *J Nucl Med Off publication Soc Nucl Med* (2012) 53(1):98–105. doi: 10.2967/jnumed.111.093831
- Pothiwala P, Jain SK, Yaturu S. Metabolic Syndrome and Cancer. *Metab Syndr Relat Disord* (2009) 7(4):279–88. doi: 10.1089/met.2008.0065
- Liu J, Nie C, Xue L, Yan Y, Liu S, Sun J, et al. Growth Hormone Receptor Disrupts Glucose Homeostasis Via Promoting and Stabilizing Retinol Binding Protein 4. *Theranostics* (2021) 11(17):8283–300. doi: 10.7150/thno.61192
- Wang Z, Embaye KS, Yang Q, Qin L, Zhang C, Liu L, et al. A Novel Metabolism-Related Signature as a Candidate Prognostic Biomarker for Hepatocellular Carcinoma. *J Hepatocell Carcinoma* (2021) 8:119–32. doi: 10.2147/JHC.S294108
- Lu Y, Wang W, Liu Z, Ma J, Zhou X, Fu W. Long Non-Coding Rna Profile Study Identifies a Metabolism-Related Signature for Colorectal Cancer. *Mol Med (Cambridge Mass)* (2021) 27(1):83. doi: 10.1186/s10020-021-00343-x
- Fan Y, Li X, Tian L, Wang J. Identification of a Metabolism-Related Signature for the Prediction of Survival in Endometrial Cancer Patients. *Front Oncol* (2021) 11:630905. doi: 10.3389/fonc.2021.630905
- Wang S, Zhang L, Yu Z, Chai K, Chen J. Identification of a Glucose Metabolism-Related Signature for Prediction of Clinical Prognosis in Clear Cell Renal Cell Carcinoma. *J Cancer* (2020) 11(17):4996–5006. doi: 10.7150/jca.45296
- Tweedie S, Braschi B, Gray K, Jones TEM, Seal RL, Yates B, et al. Genenames.Org: The Hgnc and Vgnc Resources in 2021. *Nucleic Acids Res* (2021) 49(D1):D939–D46. doi: 10.1093/nar/gkaa980
- Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, et al. Functional Genomics Reveal That the Serine Synthesis Pathway Is Essential in Breast Cancer. *Nature* (2011) 476(7360):346–50. doi: 10.1038/nature10350
- Gaujoux R, Seoighe C. A Flexible R Package for Nonnegative Matrix Factorization. *BMC Bioinf* (2010) 11:367. doi: 10.1186/1471-2105-11-367
- Hoshida Y, Brunet J-P, Tamayo P, Golub TR, Mesirov JP. Subclass Mapping: Identifying Common Subtypes in Independent Disease Data Sets. *PLoS One* (2007) 2(11):e1195. doi: 10.1371/journal.pone.0001195
- Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. Genepattern 2.0. *Nat Genet* (2006) 38(5):500–1. doi: 10.1038/ng0506-500
- Zhou J-G, Liang B, Jin S-H, Liao H-L, Du G-B, Cheng L, et al. Development and Validation of an Rna-Seq-Based Prognostic Signature in Neuroblastoma. *Front Oncol* (2019) 9:1361. doi: 10.3389/fonc.2019.01361
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for Rna-Sequencing and Microarray Studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007
- Bardou P, Mariette J, Escudié F, Djemiel C, Klopp C. Jvenn: An Interactive Venn Diagram Viewer. *BMC Bioinf* (2014) 15:293. doi: 10.1186/1471-2105-15-293
- Zhou Y-Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
- Liang B, Zhang X-X, Gu N. Virtual Screening and Network Pharmacology-Based Synergistic Mechanism Identification of Multiple Components Contained in Guanxin V Against Coronary Artery Disease. *BMC Complement Med Ther* (2020) 20(1):345. doi: 10.1186/s12906-020-03133-w

FUNDING

The study is supported by the National Natural Science Foundation of China (82072846) to BX, the Innovative Research Team of High-Level Local Universities in Shanghai (19DZ2204000) to BX, the Natural Science Foundation of Shanghai (22ZR1437300) to GY, the Interdisciplinary Program of Shanghai Jiao Tong University (project number: YG2019QNA12) to GY, the Shanghai Anticancer Association EYAS PROJECT (SACA-CY19C03) to GY, and the Fundamental Research Program Funding of Ninth People's Hospital affiliated to Shanghai Jiao Tong University School of Medicine(JYZZ006) to GY.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.909066/full#supplementary-material>

28. Hänzelmann S, Castelo R, Guinney J. Gsva: Gene Set Variation Analysis for Microarray and Rna-Seq Data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7
29. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture From Expression Data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
30. Yang R-H, Liang B, Li J-H, Pi X-B, Yu K, Xiang S-J, et al. Identification of a Novel Tumor Microenvironment-Based Prognostic Biomarker in Skin Cutaneous Melanoma. *J Cell Mol Med* (2021) 25(23):10990–1001. doi: 10.1111/jcmm.17021
31. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337
32. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A Compendium of Mutational Cancer Driver Genes. *Nat Rev Cancer* (2020) 20(10):555–72. doi: 10.1038/s41568-020-0290-x
33. Feng J, Jiang L, Li S, Tang J, Wen L. Multi-Omics Data Fusion Via a Joint Kernel Learning Model for Cancer Subtype Discovery and Essential Gene Identification. *Front Genet* (2021) 12:647141. doi: 10.3389/fgene.2021.647141
34. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
35. Wu J, Zhao W, Zhou B, Su Z, Gu X, Zhou Z, et al. Tsnadb: A Database for Tumor-Specific Neoantigens From Immunogenomics Data Analysis. *Genomics Proteomics Bioinf* (2018) 16(4):276–82. doi: 10.1016/j.gpb.2018.06.003
36. Mirchia K, Sathe AA, Walker JM, Fudym Y, Galbraith K, Viapiano MS, et al. Total Copy Number Variation as a Prognostic Factor in Adult Astrocytoma Subtypes. *Acta Neuropathologica Commun* (2019) 7(1):92. doi: 10.1186/s40478-019-0746-y
37. Zhang E, He J, Zhang H, Shan L, Wu H, Zhang M, et al. Immune-Related Gene-Based Novel Subtypes to Establish a Model Predicting the Risk of Prostate Cancer. *Front Genet* (2020) 11:595657. doi: 10.3389/fgene.2020.595657
38. Zhou J-G, Liang B, Liu J-G, Jin S-H, He S-S, Frey B, et al. Identification of 15 Lncrnas Signature for Predicting Survival Benefit of Advanced Melanoma Patients Treated With Anti-Pd-1 Monotherapy. *Cells* (2021) 10(5):977. doi: 10.3390/cells10050977
39. Yu X, Yi P, Hamilton RA, Shen H, Chen M, Foulds CE, et al. Structural Insights of Transcriptionally Active, Full-Length Androgen Receptor Coactivator Complexes. *Mol Cell* (2020) 79(5):812–23.e4. doi: 10.1016/j.molcel.2020.06.031
40. Gao X, Dai C, Huang S, Tang J, Chen G, Li J, et al. Functional Silencing of In Prostate Cancer Promotes Disease Progression. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2019) 25(4):1291–301. doi: 10.1158/1078-0432.CCR-18-2392
41. Gutjahr JC, Crawford KS, Jensen DR, Naik P, Peterson FC, Samson GPB, et al. The Dimeric Form of Cxcl12 Binds to Atypical Chemokine Receptor 1. *Sci Signaling* (2021) 14(696):eabc9012. doi: 10.1126/scisignal.abc9012
42. Sbrissa D, Semaan L, Govindarajan B, Li Y, Caruthers NJ, Stemmer PM, et al. A Novel Cross-Talk Between Cxcr4 and Pi4kiii α in Prostate Cancer Cells. *Oncogene* (2019) 38(3):332–44. doi: 10.1038/s41388-018-0448-0
43. Hong Z, Zhang W, Ding D, Huang Z, Yan Y, Cao W, et al. DNA Damage Promotes Tmprss2-Erg Oncoprotein Destruction and Prostate Cancer Suppression Via Signaling Converged by Gsk3 β and Wee1. *Mol Cell* (2020) 79(6):1008–23.e4. doi: 10.1016/j.molcel.2020.07.028
44. Zhao H, Lo Y-H, Ma L, Waltz SE, Gray JK, Hung M-C, et al. Targeting Tyrosine Phosphorylation of PcnA Inhibits Prostate Cancer Growth. *Mol Cancer Ther* (2011) 10(1):29–36. doi: 10.1158/1535-7163.MCT-10-0778
45. Zheng C-Y, Xiao W, Zhu M-X, Pan X-J, Yang Z-H, Zhou S-Y. Inhibition of Cyclooxygenase-2 by Tetramethylpyrazine and Its Effects on A549 Cell Invasion and Metastasis. *Int J Oncol* (2012) 40(6):2029–37. doi: 10.3892/ijo.2012.1375
46. Nauman G, Gray JC, Parkinson R, Levine M, Paller CJ. Systematic Review of Intravenous Ascorbate in Cancer Clinical Trials. *Antioxidants (Basel)* (2018) 7(7):89. doi: 10.3390/antiox7070089
47. Erlandsson A, Carlsson J, Lundholm M, Fält A, Andersson S-O, Andrén O, et al. M2 Macrophages and Regulatory T Cells in Lethal Prostate Cancer. *Prostate* (2019) 79(4):363–9. doi: 10.1002/pros.23742
48. Liang P, Henning SM, Schokrpur S, Wu L, Doan N, Said J, et al. Effect of Dietary Omega-3 Fatty Acids on Tumor-Associated Macrophages and Prostate Cancer Progression. *Prostate* (2016) 76(14):1293–302. doi: 10.1002/pros.23218
49. Cortesi F, Delfanti G, Grilli A, Calcinotto A, Gorini F, Pucci F, et al. Bimodal Cd40/Fas-Dependent Crosstalk Between Inkt Cells and Tumor-Associated Macrophages Impairs Prostate Cancer Progression. *Cell Rep* (2018) 22(11):3006–20. doi: 10.1016/j.celrep.2018.02.058
50. Jenkins RB, Xiao Y, Sicotte H, Decker PA, Kollmeyer TM, Hansen HM, et al. A Low-Frequency Variant at 8q24.21 Is Strongly Associated With Risk of Oligodendroglial Tumors and Astrocytomas With Idh1 or Idh2 Mutation. *Nat Genet* (2012) 44(10):1122–5. doi: 10.1038/ng.2388
51. Wilson C, Kanhere A. 8q24.21 Locus: A Paradigm to Link Non-Coding Rnas, Genome Polymorphisms and Cancer. *Int J Mol Sci* (2021) 22(3):1094. doi: 10.3390/ijms22031094
52. Liu J, Ding D, Jiang Z, Du T, Liu J, Kong Z. Long Non-Coding Rna Ccat1/Mir-148a/Pkc ζ Prevents Cell Migration of Prostate Cancer by Altering Macrophage Polarization. *Prostate* (2019) 79(1):105–12. doi: 10.1002/pros.23716
53. You Z, Liu C, Wang C, Ling Z, Wang Y, Wang Y, et al. Lncrna Ccat1 Promotes Prostate Cancer Cell Proliferation by Interacting With Ddx5 and Mir-28-5p. *Mol Cancer Ther* (2019) 18(12):2469–79. doi: 10.1158/1535-7163.MCT-19-0095
54. Liu H-T, Fang L, Cheng Y-X, Sun Q. Lncrna Pvt1 Regulates Prostate Cancer Cell Growth by Inducing the Methylation of Mir-146a. *Cancer Med* (2016) 5(12):3512–9. doi: 10.1002/cam4.900
55. Quigley DA, Dang HX, Zhao SG, Lloyd P, Aggarwal R, Alumkal JJ, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. *Cell* (2018) 174(3):758–69.e9. doi: 10.1016/j.cell.2018.06.039
56. Hua JT, Ahmed M, Guo H, Zhang Y, Chen S, Soares F, et al. Risk Snp-Mediated Promoter-Enhancer Switching Drives Prostate Cancer Through Lncrna Pcat19. *Cell* (2018) 174(3):564–75.e18. doi: 10.1016/j.cell.2018.06.014
57. Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, et al. Insertion-And-Deletion-Derived Tumour-Specific Neoantigens and the Immunogenic Phenotype: A Pan-Cancer Analysis. *Lancet Oncol* (2017) 18(8):1009–21. doi: 10.1016/S1470-2045(17)30516-8
58. Antonarakis ES, Isaacsson Velho P, Fu W, Wang H, Agarwal N, Sacristan Santos V, et al. -Altered Prostate Cancer: Clinical Features and Therapeutic Outcomes to Standard Systemic Therapies, Poly (Adp-Ribose) Polymerase Inhibitors, and Pd-1 Inhibitors. *JCO Precis Oncol* (2020) 4:370–81. doi: 10.1200/po.19.00399
59. Collins JM, Redman JM, Gulley JL. Combining Vaccines and Immune Checkpoint Inhibitors to Prime, Expand, and Facilitate Effective Tumor Immunotherapy. *Expert Rev Vaccines* (2018) 17(8):697–705. doi: 10.1080/14760584.2018.1506332
60. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, et al. Sipuleucel-T Immunotherapy for Castration-Resistant Prostate Cancer. *New Engl J Med* (2010) 363(5):411–22. doi: 10.1056/NEJMoa1001294

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

Reviewer DH declared a shared parent affiliation with authors GY, MZ, XG, JW, SS, YL, QY, TJ, and BX to the handling editor at the time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yu, Liang, Yin, Zhan, Gu, Wang, Song, Liu, Yang, Ji and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.