Identification of immune-based prostate cancer subtypes using mRNA expression

Jukun Song (songjukun@163.com)
Guizhou provinicial people"s hospital https://orcid.org/0000-0003-2542-9340

Song He
School Of Medicine, Guizhou University, Guizhou, China.

Wei Wang
Department of Urology, Guizhou Provincial People's Hospital

Jiaming Su
Guizhou Provincial People's Hospital

Dongbo Yuan
Department of Urology, Guizhou Provincial People's Hospital

Weihong Chen
Department of Urology, Guizhou Provincial People's Hospital

Jianguo Zhu
Department of Urology, Guizhou Provincial People's Hospital

Research

Keywords: Prostate cancer, Tumor-infiltrating immune cells, CIBERSORT algorithm, Clinical application, Immunotherapy

Posted Date: February 10th, 2020

DOI: https://doi.org/10.21203/rs.2.23038/v1

Version of Record: A version of this preprint was published at Bioscience Reports on January 1st, 2021.
See the published version at https://doi.org/10.1042/BSR20201533.

Abstract

Background

Immune infiltration of Prostate cancer (PCa) was highly related to clinical outcomes. However, previous works failed to elucidate the diversity of different immune cell types that make up the function of the immune response system. The aim of the study was to uncover the composition of TIICs in PCa utilizing the CIBERSORT algorithm and further reveal the molecular characteristics of PCa subtypes.

Method

In the present work, we employed the CIBERSORT method to evaluate the relative proportions of immune cell profiling in PCa and adjacent samples, normal samples. We analyzed the correlation between immune cell infiltration and clinical information. The tumor-infiltrating immune cells of the TCGA PCa cohort were analyzed for the first time. The fractions of 22 immune cell types were imputed to determine the correlation between each immune cell subpopulation and clinical feature. Three types of molecular classification were identified via R-package of “CancerSubtypes”. The functional enrichment was analyzed in each subtype. The submap and TIDE algorithm were used to predict the clinical response to immune checkpoint blockade, and GDSC was employed to screen chemotherapeutic targets for the potential treatment of PCa.

Results

In current work, we utilized the CIBERSORT algorithm to assess the relative proportions of immune cell profiling in PCa and adjacent samples, normal samples. We investigated the correlation between immune cell infiltration and clinical data. The tumor-infiltrating immune cells in the TCGA PCa cohort were analyzed. The 22 immune cells were also calculated to determine the correlation between each immune cell subpopulation and survival and response to chemotherapy. Three types of molecular classification were identified. Each subtype has specific molecular and clinical characteristics. Meanwhile, Cluster I is defined as advanced PCa, and is more likely to respond to immunotherapy.

Conclusions

Our results demonstrated that differences in immune response may be important drivers of PCa progression and response to treatment. The deconvolution algorithm of gene expression microarray data by CIBERSOFT provides useful information about the immune cell composition of PCa patients. In addition, we have found a subtype of immunopositive PCa subtype and will help to explore the reasons for the poor effect of PCa on immunotherapy, and it is expected that immunotherapy will be used to guide the individualized management and treatment of PCa patients.

Background

Prostate cancer (PCa) is the most common malignancy in Europe and the United States. Among American male cancer patients, patients who died of prostate cancer ranked second, second only to breast cancer [1]. In 2019, the American Cancer Society announced 174,650 new cases of prostate cancer, accounting for 20% of new male cancer cases, ranking first, with 31,620 deaths from prostate cancer, accounting for 10% of total cancer deaths [2]. Prostate cancer is the first major tumor type in 28 European countries, and the second most prominent type in seven other countries[3]. The ethnic differences in the incidence of prostate cancer are very obvious. The incidence and mortality rate of PCa in China is lower than that in Western countries such as Europe and the United States. However, with the development of society and changes in people's lifestyles, prostate cancer has become a common tumor in the male urinary tract, and its incidence has increased year by year [4]. In addition, PCa is a heterogeneous disease that can vary greatly even within the same tumor[5]. Significant differences in incidence and morbidity may be due to genomic instability and changes associated with various PCa risk factors. The treatment in early PCa has achieved satisfactory effects through androgen deprivation therapy, but ultimately it was inevitable to develop hormone-dependent PCa, which caused obvious clinical symptoms [6]. However, the phenotype of cancer is defined not only by the intrinsic activity of tumor cells but also by immune cells recruited and attracted in the microenvironment associated with the tumor. At present, the role of immune cells in the tumor-associated microenvironment during tumor development has not been known, especially in PCa development.

A growing body of evidence demonstrated that tumor-infiltrating immune cells (TIICs), including B cells, T cells, dendritic cells, macrophages, neutrophils, monocytes, and polar cells, might control malignant growth in the tumor. TIICs are important components of the tumor microenvironment and can alter the immune status of tumors. The impact of TIICs on malignancy advancement has been to a great extent documented [7–10]. In the present work, a newly developed deconvolution algorithm, CIBERSORT, was used to define 22 TIICs subsets of immune responses to examine correlations with molecular subsets and clinical features. Our findings also revealed distinct immune phenotypes for molecular PCa subclasses. In addition, the present investigation gives a novel understanding of immunotherapy methodology for PCa.

Materials And Methods

Gene expression datasets

Publicly available gene expression profiles from TCGA and GTEX databases were enrolled. The human healthy prostate tissues were available from the GTEX (https://gtexportal.org/home/), while the adjacent tumor samples and tumor samples were obtained from The Cancer Genome Atlas (TCGA) site (https://cancergenome.nih.gov/). The RNA-seq profiles (FPKM values) and phenotype data were downloaded. Firstly, the voom (variance modeling at the observational level) method was used to convert the mRNA sequencing data into similar results for the microarray data [11]. We then annotated the gene ID and normalized the mRNA sequencing results using the "limma" package of R software to average the repeated gene data and remove the unavailable data[12]. We manually organized the expression matrix

and clinical characteristics of each patient. Patients with fully clinical pathology data and survival time of more than 30 days were included in the study.

Inference Of TIICs Using The CIBERSOFT Algorithm

The CIBERSOFT is an analytical tool which accurately infers the relative levels of human immune cell types within a complex gene expression profiles. CIBERSOFT uses the characteristics of 547 marker genes to characterize and quantify the relative scores for each immune cell subtype. To enhance the robustness of the results, the CIBERSOFT algorithm uses Monte Carlo sampling to derive the deconvoluted P-value for each sample. The standardized processed dataset of gene expression is uploaded to the CIBERSOFT website (https://cibersort.stanford.edu/index.php), which runs utilizing 1000 aligned default signature matrices [13]. After using the CIBERSORT program, the distribution of the LM22 subtypes of TIICs together with the results of the correlation coefficients, P-value and root mean square error (RMSE), can be used to assess the accuracy of the results in each sample. A P value of ≤ 0.05 reflects the statistical significance of the deconvolution result of each sample on all subsets of cells and is useful for screening results with lower precision. Finally, 64 normal samples, 32 adjacent tumor samples, and 351 tumor samples were selected for further analysis with the cut-off of P value less than 0.05.

We inferred total T cell fraction as a sum of CD8 + T cells, CD4 + naïve T cells, CD4 + memory resting T cells, CD4 + memory activated T cells, follicular helper T cells, regulatory T cells (Tregs) and T cells gamma delta fractions. Total macrophage cell fraction was calculated as a sum of M0, M1 and M2 macrophage fractions. Total B cell fraction was imputed as a sum of B cell memory and B cells naïve.

Identification Of Prognostic LM22 Immune Cell Subset

The prognostic LM22 immune cell phenotypes were identified. Firstly, the univariate Cox analysis and Kaplan-Meier survival analysis was performed to screen prognostic immune cell type. Then, the multivariate Cox regression analysis was used to further validate prognostic 22 human immune cell phenotypes as prognosis factors.

Unsupervised Clustering Analysis

A consensus cluster algorithm was applied to determine the number of clusters across the tumor samples. We used the SNFCC + method in the “CancerSubtypes” R-package to identify cancer subtype for genomic data[14]. The performance of these clustering methods was evaluated by three common methods: (1) Log-rank test of Kaplan-Meier curves to evaluate the importance of survival differences between subtypes; (2) Average contour width (ASW) method of measuring cluster consistency for

evaluating whether samples within a subtype or subtypes are more similar; (3) Clustering heat maps to visually visualize sample clusters by separating color patches from each other.

The screen of differentially expressed genes (DEGs) and differential immune cell types in each subclass

To uncover the potential subtype–specific immune related genes and TIIC models in each cluster, differentially expressed genes (DEGs) and specific immune cell types were identified using the R package limma among 547 marker genes, which implements an empirical Bayesian approach to estimate gene expression changes using moderated t-tests. DEGs and differential immune cell types were determined by significance criteria (adjusted P-value < 0.05)[12].

Pathways And Biological Functions Differentially Enriched PCa Subtypes

Functional enrichment analysis using the “Cluserprofiler” of R package was conducted on DEGs among PCa subtypes[15]. Gene Ontology (GO) Biological Process term and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms were identified with a cutoff of P < 0.05. Gene Set Enrichment Analysis (GSEA) was conducted to unveil an overall pathway of gene-set activity score for each sample[16]. The Gene sets using the c2/c5 curated signatures were downloaded from the Molecular Signature Database (MSigDB) of Broad Institute. The significant enrichment pathway was identified based on the adjusted P-value < 0.05.

Mutation Analysis

In the genetic and epigenetic analysis, mutation data in the MAF of 301 PCa patients were used. The R package "maftools" was used to display the mutation profile of each subtype[17]. The maftools was also used to impute the mutation rate of each gene and also identified significant mutant genes in different subtypes (P < 0.05).

Prediction For Chemo/immunotherapeutic Response

Tumor immune dysfunction and exclusion (TIDE) algorithms[18] and subclass mapping[19] are used to predict clinical response to immune checkpoints. The chemotherapeutic response of each sample was also predicted based on the largest publicly available pharmacogenomics database [Pharmaceutical Sensitivity Genomics in Cancer (GDSC), https://www.cancerrxgene.org/][20]. The prediction procedure was performed by the R software package "pRRophetic", where the half-maximal inhibitory concentration

(IC50) of the samples was demonstrated by ridge regression and the prediction accuracy was assessed by 10-fold cross-validation based on the GDSC training set[21].

Statistical analysis

All statistical tests were analyzed by R (3.5.2) utilizing a $\chi 2$ or Fisher's exact test for categorical data when suitable. As for continuous data, a Wilcoxon test (Mann-Whitney test) for two groups, the Kruskal-Wallis test for more than two groups [22]; As for survival data, Kaplan-Meier curve[23] and Cox regression[24] was conducted to screen prognostic immune cell subclasses. Survival analysis was performed using the R package "survival". Fisher's independence exact test is used to statistically classify the relationship between clinical information and defined subtypes. For all statistical analyses, a P value of less than 0.05 was considered statistically significant.

Results

Inference of immune infiltration cells in TME

The landscape of TME cell infiltration models was inferred by the CIBERSOFT algorithm. Figure 1A,B summarizes the findings obtained from the 64 normal samples, 32 adjacent tumor samples, and 351 tumor samples. Twenty-one types of TIICs were detected in patients, while naïve CD4 + T cells did not appear in any samples, the findings were in accordance with the previous study[25]. The Correlation matrix of all 21 immune cell densities in the TCGA cohort was also shown in Fig. 1C,D. The results in current works found that the proportion of TME cells in PCa was significantly different among the three groups. In general, the infiltration of TME cells differed among these samples, with the highest proportion of LM22 immune cells found in normal samples, followed by adjacent tumor samples, and finally tumor samples. The mutual interaction relationship between immune cells in normal and adjacent tumor samples is evident compared to tumor samples (Fig. 2A,B,C,D). Compare to normal tissues, the fraction of total T cells and total B cells were higher in adjacent tumor tissues and tumor tissues. Total Macrophage was mainly observed in the tumor tissue (Fig. 3A,B,C). B cell memory was devoid in the adjacent tumor samples and tumor samples, Macrophages M1 and M2 were increased in the tumor samples, T cells gamma delta and T cells regulatory (Tregs) was found in the adjacent samples and tumor samples. We can infer from the results that the immune system is inhibited during tumor development.

We found that B cell memory, Macrophages M2, Mast cells activated, Mast cells resting, Monocytes, NK cells resting, NK cells activated, T cells CD8 and T cell follicular helper were mainly enriched in the normal tissues, while B cell naive and Dendritic cells activated, Dendritic cells resting, Macrophages M1, T cells CD4 memory activated, T cells CD4 memory resting, Neutrophils were significantly founded in the adjacent tissues. In the tumor tissues, B cell naïve, Dendritic cells activated, Dendritic cells resting, Macrophages M0, Macrophages M1, T cells CD4 memory activated, T cells CD4 memory resting and T cells regulatory (Tregs) were mainly activated (Fig. 4). Therefore, these results demonstrated that aberrant

immune infiltration and it's heterogeneous in PCa as a tightly regulated process might play important roles in the development of tumor, it has important clinical meanings.

Prognostic Subsets Of Immune Cells

The univariate Cox regression analysis was performed to identify the prognostic subtypes of TIICs in PCa, and the results were shown that B cell memory was significantly correlated with RFS with the cut-off of the P-value less than 0.05 (Table 1, Fig. 5A). In the multivariate Cox regression analysis, B cell memory was still obviously related to RFS (Table 2, Fig. 5B). Then, we performed a Kaplan-Meier curve plot and log-rank test for the above-mentioned immune cell subsets, and the results are shown in Fig. 5C,D. B cell memory and T cells regulatory (Tregs) were significantly associated with RFS with PCa patients.

Table 1
The univariate Cox regression analysis was performed to identify the prognostic subtypes of TIICs in PCa.

Immune cell	HR	HR.95L	HR.95H	pvalue
B cells naive	0.663123	0.000441	996.5155	0.912357
B cells memory	2.98E + 34	2.19E + 13	4.06E + 55	0.001388
Plasma cells	0.001334	1.20E-08	148.7424	0.264269
T cells CD8	2.075994	0.005774	746.3596	0.80779
T cells CD4 memory resting	0.253485	0.000796	80.75887	0.640723
T cells CD4 memory activated	0.003227	2.03E-12	5119968	0.595629
T cells follicular helper	0.000369	7.62E-10	178.984	0.236668
T cells regulatory (Tregs)	79.00902	2.48E-08	2.52E + 11	0.695531
T cells gamma delta	0.000811	3.25E-22	2.02E + 15	0.741926
NK cells resting	2831.234	1.18E-30	6.77E + 36	0.839372
NK cells activated	2.001437	2.76E-07	14491266	0.931388
Monocytes	3357.781	0.000786	1.43E + 10	0.297283
Macrophages M0	11.70166	0.01414	9683.931	0.473023
Macrophages M1	2762.816	0.007955	9.6E + 08	0.223474
Macrophages M2	70.26358	0.016501	299185.2	0.318604
Dendritic cells resting	0.056424	3.43E-06	927.6329	0.561617
Dendritic cells activated	1.11E-14	3.10E-31	399.1275	0.098531
Mast cells resting	2.98651	0.001924	4635.156	0.770392
Mast cells activated	4502.05	1.40E-19	1.45E + 26	0.75037
Eosinophils	8.19E-34	5.81E-99	1.16E + 32	0.319553
Neutrophils	2.02E-39	3.99E-89	1.02E + 11	0.127062

Table 2
The multivariate Cox regression analysis was conducted to screen the prognostic subtypes of TIICs in PCa.

Immune cell	HR	HR.95L	HR.95H	pvalue
B cells naive	6.67E + 39	7.90E-12	5.63E + 90	0.125353
B cells memory	6.60E + 69	9.74E + 12	4.48E + 126	0.016044
Plasma cells	2.19E + 35	1.15E-16	4.18E + 86	0.176776
T cells CD8	3.96E + 37	7.90E-14	1.99E + 88	0.146099
T cells CD4 memory resting	6.14E + 36	4.59E-15	8.20E + 87	0.158454
T cells CD4 memory activated	2.66E + 37	5.70E-16	1.24E + 90	0.163714
T cells follicular helper	1.92E + 34	3.94E-18	9.33E + 85	0.193608
T cells regulatory (Tregs)	2.48E + 38	3.96E-14	1.55E + 90	0.146276
T cells gamma delta	8.99E + 38	8.48E-17	9.52E + 93	0.165286
NK cells resting	4.12E + 53	3.86E-18	4.40E + 124	0.139016
NK cells activated	2.11E + 38	4.75E-13	9.35E + 88	0.138054
Monocytes	2.04E + 44	3.07E-08	1.36E + 96	0.093777
Macrophages M0	1.19E + 38	2.38E-13	5.92E + 88	0.141039
Macrophages M1	4.64E + 40	1.24E-10	1.74E + 91	0.115014
Macrophages M2	2.82E + 38	3.46E-13	2.29E + 89	0.138805
Dendritic cells resting	1.98E + 38	1.84E-13	2.13E + 89	0.141324
Dendritic cells activated	3.37E + 20	4.47E-34	2.54E + 74	0.455198
Mast cells resting	3.10E + 38	2.19E-13	4.37E + 89	0.140241
Mast cells activated	2.80E + 55	0.00022	3.57E + 114	0.065967
Eosinophils	4.63E-10	4.73E-95	4.52E + 75	0.82956
Neutrophils	NA	NA	NA	NA

Patterns Of Immune Cell Infiltration In Molecular PCa Subclasses

The molecular classification of PCa was identified by unsupervised consensus clustering in all tumor samples using“CancerSubtypes” R-package. The optimal number of clusters is determined by the K

value. After evaluating the relative changes in the area under the cumulative distribution function (CDF) curve and the consensus heat map, a three-cluster solution (K = 3) with no significant increase in area under the CDF curve was selected (Fig. 6). This result revealed 88 patients (38%) in Cluster I, 92 patients (48%) in Cluster II, and 133 patients (64%) in Cluster III for the PCa cohort. The consensus matrix heat map demonstrated that cluster I and cluster II, and clusters III appear well in the individualized clusters in Fig. 7(A, B, C). Patients classified as Cluster III had a good prognosis compared to Cluster I and Cluster II ($P < 0.001$, log-rank test).

Differentially expressed genes and immune cell patterns of Cluster I, Cluster II and Cluster III subtypes

The molecular subtypes of PCa cause differences in clusters I, II, and III subtypes, which exhibit activation of specific signaling pathways and different prognosis. The Kruskal-Wallis test was performed to identify subtype-specific genes/LM22 immune cells in each subtype. In terms of subtype-specific immune cells in PCa, Cluster I and III is defined by high levels of Dendritic cells resting compared with Cluster II. Cluster II was enriched by Macrophages M0 and NK cells activated; While Cluster III was defined by the high level of Dendritic cells activated, Dendritic cells resting, T cells CD4 memory activated and T cells regulatory (Tregs) (Fig. 8).

To uncover molecular differences among PCa molecular subtypes and identify subtype-specific biomarkers, unpaired Student's t-test was used to identify differentially expressed genes that were significantly associated with each subtype. Gene differentially expressed analysis was conducted with the cut-off point of FDR < 0.05. In subgroup I, a total of 294 immune related mRNAs (227 up-regulated genes and 67 down-regulated genes) were differentially expressed genes (DEGs) compared to other groups; In subgroup II, a total of 292 DEGs (76 up-regulated genes and 216 down-regulated genes) compared other groups; in subgroup III, a total of 162 DEGs (162 up-regulated genes and 95 down-regulated genes) were observed compared to other groups (Fig. 9).

In terms of clinical features among three clusters, Cluster I has a Higher Gleason score and PSA level compared to Cluster II and III. However, the PSA value and Age has no difference between the three Clusters (Fig. 10A,B,C). The heatmap illustrates the association of different clinical characters among the three subgroups (Fig. 10D). Statistical significance was performed by the Kruskal-Wallis test.

Identification of differentially expressed genes and enriched Gene Ontology and pathway in subtypes

Functional enrichment analysis for DEGs in CI vs CII, III, CII vs CI, III and C III vs CI, II was performed. For CI vs CII, III,a total of 517 GO terms of biological process, 25 GO terms of cellular component, and 39 GO terms of molecular function were identified with the cutoff point of adjust P-value < 0.05 (Supplementary

Table 1). In subgroup II compared to other groups, 476 GO terms of biological process, 25 GO terms of cellular component, and 45 GO terms of molecular function were identified. In subgroup III compared to other groups, a total of 425 GO terms of biological process, 13 GO terms of cellular component, and 18 GO terms of molecular function were identified. Top GO terms included cytokine activity, immune/inflammatory response, and chemokine activity (Fig. 11A). All Clusters were enriched in cytokine activity, chemokine activity and carbohydrate binding. Cluster I and II were associated with lipopeptide binding, while Cluster III was related to receptor ligand activity and G protein – coupled purinergic nucleotide receptor activity, G protein – coupled nucleotide receptor activity (Fig. 12A).

Furthermore, all pathways generated by the Kyoto Genomics and Genomics Encyclopedia (KEGG) analysis are related to immune responses (Fig. 11B). Cluster I was associated with the Primary immunodeficiency, Th1 and Th2 cell differentiation and NF – kappa B signaling pathway, Cluster II was enriched for Primary immunodeficiency, JAK – STAT signaling pathway and Natural killer cell mediated cytotoxicity, while Cluster III was related to Th17 cell differentiation, Cell adhesion molecules (CAMs) and TNF signaling pathway (Fig. 12B).

To identify gene sets enriched in each subtype, we then performed GSEA analysis. GSEA analysis reveals distinct enriched gene sets between subtypes. The number of enriched pathways progressively increased from Cluster I through Cluster II. We subsequently selected representative gene sets for CI-CIII to build a pathway heatmap, which reveals distinct gene sets enriched in each subtype. Cluster I was associated with SMID BREAST CANCER ERBB2 and SHEDDEN LUNG CANCER POOR SURVIVAL, Cluster II was enriched for RUTELLA_RESPONSE_TO_HGF_VS_CSF2RB_AND_IL4_DN and VILIMAS_NOTCH1_TARGETS, while Cluster III was related to BASSO CD40 SIGNALING and VERHAAK AML WITH NPM1 MUTATED (Fig. 13A,B,C,D). All these similarities are in good agreement with the molecular and clinical characteristics of the three subtypes identified in our study, confirming the correctness of the features we found on these three subtypes.

Subclass-associated Gene Mutations

This study also investigated the association between the three subtypes and count of somatic mutation. Gene mutation profiles of these highly mutated genes are shown in Fig. 14A,B,C,D,E,F. SPOP showed a higher mutation rate in Cluster I, TP53 exhibited a higher mutation rate in Cluster II, while in the Cluster III, CACNA1E and TTN, TP53 exhibited a higher mutation rate.

Sensitivity to Immuno/Chemotherapies among Cluster I, II and III Subtype

Although immunological checkpoint drugs have not been approved for use as a conventional drug for PCa, we have used the TIDE algorithm to predict the likelihood of responding to immunotherapy, and it has demonstrated that cluster III was more likely better to the response to immunotherapy than clustering

II and III ($P < 0.05$). As for the TIDE prediction, a subclass mapping method was employed to compare the expression profiles of the three PCa subtypes with another published data set containing 47 melanoma patients who responded to immunotherapy. The results showed that Cluster I was more promising for CTLA4 treatment (Bonferroni correction $P < 0.05$) (Fig. 15A).

Given that chemotherapy is currently a common treatment for prostate cancer, we sought to evaluate the response of three subtypes to common chemotherapy drugs. Therefore, we trained the prediction model on the GDSC cell line dataset by ridge regression and evaluated the satisfactory prediction accuracy by 10-fold cross-validation. We estimated the IC 50 for each sample in the TCGA dataset based on the predictive model of these two chemo drugs. For several common chemicals, we can observe that Cluster I and Cluster II, Cluster III have significant differences in estimated IC 50, of which Cluster I may be more sensitive to commonly used chemotherapy (AKT.inhibitor.VIII $P < 0.001$, ATRA $P = < 0.001$) (Fig. 15B).

Discussion

Prostate cancer is a major disease that threatens human health. Prostate cancer (PCa) is the most common malignancy in Western countries[26]. Cancer has gradually been recognized as an adaptive and complex system, and it is increasingly difficult to achieve the desired therapeutic effect by using too many single-targeted drugs. To overcome or prevent drug resistance, immunotherapy has become promising therapeutic targets for cancer treatment. At present, immunotherapy has achieved satisfactory results in cancer treatment, such as malignant melanoma, but immunotherapy is not effective in the treatment of prostate cancer[27, 28]. Alternatively, TIICs may be effective targets for drugs to improve clinical outcomes. With the increasing focus on immune checkpoint therapy, the distribution of TIIC has become a major research spot hot. Previous studies estimated that the degree of TIIC is mainly based on immunohistochemistry (IHC), but the markers of immunohistochemistry are not precise, and multiple TIICs may express the same markers on the membrane, which leads to the measurement of TIIC density in PCa tumor tissue accurate[29]. Therefore, this study aims to predict the immune cell infiltration using CIBERSOFT algorithm in the TCGA PCa cohort and to explore the evaluation of therapeutic efficacy of immunotherapy in different subtypes, and to find a specific subtype that is effective for immunotherapy.

In the current work, the CIBERSOFT algorithm was used to assess the infiltration of immune cells among normal tissues and adjacent tumor tissues, PCa tumor tissues, and the results showed that there were considerable differences in immune cell components within and between groups. Our work also revealed details of the subpopulation of LM22 immune cells in PCa. Macrophages and natural killer (NK) cells have been detected in PCa[30, 31]. In addition, NK cells and macrophages can be used in androgen deprivation therapy, and that a large number of NK cells can reduce the risk of tumor progression, but a high proportion of macrophages can increase the risk of biochemical recurrence[32]. It is urgent that new and increasingly successful prognostic markers should be added to the current quality marks to improve expectation precision.

Our work also uncovers the detail of infiltration of LM22 immune cell subsets in PCa that the proportions of total T cells infiltration fraction account for more than 39%, in which 20% is T cells CD4 memory resting, 13% is T cells CD8, T cells follicular helper make up 3%, and T cells CD4 memory activated and T cells regulatory (Tregs), T cells gamma delta were less than 1%. In addition, total Macrophages cells as the second proportion (13%) (Table 1). Compared to normal and adjacent tumor tissues, the proportions of total T cells, total B cells, Dendritic cells activated, Macrophages M0, Macrophages M1, NK cells activated and T cells CD4 memory produced statistical significance ($P < 0.05$). The Dendritic cells resting and Neutrophils were increased in adjacent tumor tissues. In contrast, B cell memory, Macrophages M2, Mast cells activated, Mast cells resting, NK cells activated, NK cells resting, Plasma cells and T cell follicular helper were mainly enriched in the normal tissues. The mechanisms behind the activation of NK cells and Macrophages, Dendritic cells activated, deactivation of B cells and Plasma cells in PCa remain unclear. Macrophages cells and NK cells have been detected in PCa[33, 34]. Philippe et al. reported that patients with androgen deprivation therapy can use natural killer (NK) cells and macrophages, and a large number of NK cells can reduce the risk of tumor progression, but a high proportion of macrophages can increase the risk of biochemical recurrence[32].

The prognostic importance of immune cell infiltration for various solid tumor types has been determined[35–37]. In univariate and multivariate Cox analysis, only B cell memory was significantly correlated with RFS. Patients whose B cell memory density is higher had a shorter RFS time. Meanwhile, the KM curve of B cell memory showed a similar trend. Several studies reported that clinical data show tumor-infiltrating B cells are a positive prognostic factor, such as breast cancer[38], ovarian cancer[39], non-small cell lung cancer[40], and other cancers[41, 42]. Spear et al demonstrated that B cells are active members of the tumor microenvironment and can produce immune-stimulating factors that may support adaptive anti-tumor immune responses in pancreatic ductal adenocarcinoma[43].

PCa can be reliably divided into three subtypes by the SNFCC + method. The results showed that the three subtypes were significantly associated with patient survival. Patients classified as class II have a better prognosis than class II and class III. Compared to each cluster of LM22 immune cells in HNC TME, cluster I was defined by high levels of macrophage infiltration, and cluster II was enriched by B cell infiltration and T cell infiltration. Although cluster III is defined by high levels of mast cell infiltration, neutrophil infiltration, and NK cell infiltration. Each cluster has its own feature-rich terminology compared to each DEG cluster. Therefore, these three subtypes differ in RFS and molecular characteristics. The submap and TIDE analysis suggested that Cluster I was more promising for CTLA4 treatment. Using the GDSC database, we imputed that Cluster I could be more sensitive to commonly used chemotherapies than Cluster II and III. The above discussion implicates that Cluster I may benefit from the combination of chemotherapy and immunotherapy, and Cluster I may be more sensitive to commonly used chemotherapy (AKT.inhibitor.VIII $P < 0.001$, ATRA $P = < 0.001$).

Conclusions

Our analysis of the LM22 immune cell subset using the CIBERSORT deconvolution algorithm provides complete information on the pattern of PCa immune cells. Our findings also found an important role in predicting clinical outcomes. In the current work, this comprehensive assessment of the LM22 immune cell infiltration model in TME opens up the possibility of how tumors respond to immunotherapy and may help clinicians explore the development of new drugs.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Data availability could be obtained from TCGA website.

Competing interests

The authors declare that they have no competing interests.

Author contribution statement:

J.K.S; D.B.Y; J.G.Z; W.H.C wrote the main manuscript text;

J.K.S; J.G.L;J.M.S prepared Figures 1-15;

H.S; J.G.Z;W.W contributed to data analysis;

All authors reviewed the manuscript.

Acknowledgment

None.

Funding statement

This study was supported by the NSFC (Natural Science Foundation of China) (grant no. 81873608)□ Mechanism of PGE2 activating ureteral smooth muscle cAMP signaling pathway involved in ureteral drainage; The Science and Technology Plan Project of Guizhou Province in 2019 (grant no. [2019]2799): Miao Yao Ning Tetai promotes stone discharge by relaxing ureteral smooth muscle with PGE2/cAMP signaling pathway; The High-level innovative talent project of Guizhou Province in 2018 (grant no. [2018]5639): Guizhou high-level innovative talent project (100 levels) and The Science and Technology

Plan Project of Guiyang in 2019 (grant no. [[2019]2-15]: Application of Artificial Intelligence in the Background of Big Data in Accurate Diagnosis and Treatment of Prostate Cancer.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018; 68: 394-424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019; 69: 7-34.
3. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019; 144: 1941-53.
4. Liu X, Yu C, Bi Y, Zhang ZJ. Trends and age-period-cohort effect on incidence and mortality of prostate cancer from 1990 to 2017 in China. Public Health. 2019; 172: 70-80.
5. Carm KT, Hoff AM, Bakken AC, Axcrona U, Axcrona K, Lothe RA, Skotheim RI, Lovf M. Interfocal heterogeneity challenges the clinical usefulness of molecular classification of primary prostate cancer. Sci Rep. 2019; 9: 13579.
6. Kgatle MM, Kalla AA, Islam MM, Sathekge M, Moorad R. Prostate Cancer: Epigenetic Alterations, Risk Factors, and Therapy. Prostate Cancer. 2016; 2016: 5653862.
7. Shibutani M, Maeda K, Nagahara H, Fukuoka T, Iseki Y, Matsutani S, Kashiwagi S, Tanaka H, Hirakawa K, Ohira M. Tumor-infiltrating Lymphocytes Predict the Chemotherapeutic Outcomes in Patients with Stage IV Colorectal Cancer. In Vivo. 2018; 32: 151-8.
8. Zhang S, Zhang E, Long J, Hu Z, Peng J, Liu L, Tang F, Li L, Ouyang Y, Zeng Z. Immune infiltration in renal cell carcinoma. Cancer Sci. 2019; 110: 1564-72.
9. Aponte-Lopez A, Fuentes-Panana EM, Cortes-Munoz D, Munoz-Cruz S. Mast Cell, the Neglected Member of the Tumor Microenvironment: Role in Breast Cancer. J Immunol Res. 2018; 2018: 2584243.
10. Cai DL, Jin LP. Immune Cell Population in Ovarian Tumor Microenvironment. J Cancer. 2017; 8: 2915-23.
11. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014; 15: R29.
12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43: e47.
13. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12: 453-7.
14. Xu T, Le TD, Liu L, Su N, Wang R, Sun B, Colaprico A, Bontempi G, Li J. CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization. Bioinformatics. 2017; 33: 3131-3.

15. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16: 284-7.
16. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14: 7.
17. Mayakonda A, Koeffler HP. Maftools: Efficient analysis, visualization and summarization of MAF files from large-scale cohort based cancer studies. BioRxiv. 2016: 052662.
18. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018; 24: 1550-8.
19. Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP. Subclass mapping: identifying common subtypes in independent disease data sets. PLoS One. 2007; 2: e1195.
20. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013; 41: D955-61.
21. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. Genome Biol. 2014; 15: R47.
22. Hazra A, Gogtay N. Biostatistics Series Module 3: Comparing Groups: Numerical Variables. Indian J Dermatol. 2016; 61: 251-60.
23. Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). BMJ. 1998; 317: 1572.
24. Fox J. Cox proportional-hazards regression for survival data. An R and S-PLUS companion to applied regression. 2002; 2002.
25. Meng J, Liu Y, Guan S, Fan S, Zhou J, Zhang M, Liang C. The establishment of immune infiltration based novel recurrence predicting nomogram in prostate cancer. Cancer Med. 2019; 8: 5202-13.
26. Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, Gavin A, Visser O, Bray F. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer. 2018; 103: 356-87.
27. Kwiatkowska D, Kluska P, Reich A. Beyond PD-1 Immunotherapy in Malignant Melanoma. Dermatol Ther (Heidelb). 2019; 9: 243-57.
28. Koller KM, Wang W, Schell TD, Cozza EM, Kokolus KM, Neves RI, Mackley HB, Pameijer C, Leung A, Anderson B, Mallon CA, Robertson G, Drabick JJ. Malignant melanoma-The cradle of anti-neoplastic immunotherapy. Crit Rev Oncol Hematol. 2016; 106: 25-54.
29. Chevrier S, Levine JH, Zanotelli VRT, Silina K, Schulz D, Bacac M, Ries CH, Ailles L, Jewett MAS, Moch H, van den Broek M, Beisel C, Stadler MB, et al. An Immune Atlas of Clear Cell Renal Cell Carcinoma. Cell. 2017; 169: 736-49 e18.
30. Zhang Q, Xia J, Wang Y, Zhang J, Ji C, Cong R, Wang Y, Song N. Tumor infiltrating M2 macrophages could predict biochemical recurrence of localized prostate cancer after radical prostatectomy. Exp

Cell Res. 2019; 384: 111588.

31. Hood SP, Foulds GA, Imrie H, Reeder S, McArdle SEB, Khan M, Pockley AG. Phenotype and Function of Activated Natural Killer Cells From Patients With Prostate Cancer: Patient-Dependent Responses to Priming and IL-2 Activation. Front Immunol. 2018; 9: 3169.
32. Gannon PO, Poisson AO, Delvoye N, Lapointe R, Mes-Masson AM, Saad F. Characterization of the intra-prostatic immune cell infiltration in androgen-deprived prostate cancer patients. J Immunol Methods. 2009; 348: 9-17.
33. Chen X, Fu E, Lou H, Mao X, Yan B, Tong F, Sun J, Wei L. IL-6 induced M1 type macrophage polarization increases radiosensitivity in HPV positive head and neck cancer. Cancer Lett. 2019.
34. Friedman J, Padget M, Lee J, Schlom J, Hodge J, Allen C. Direct and antibody-dependent cell-mediated cytotoxicity of head and neck squamous cell carcinoma cells by high-affinity natural killer cells. Oral Oncol. 2019; 90: 38-44.
35. Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, Bin J, Liao Y, Liao W. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer. Cancer Immunol Immunother. 2019; 68: 433-42.
36. Miksch RC, Schoenberg MB, Weniger M, Bosch F, Ormanns S, Mayer B, Werner J, Bazhin AV, D'Haese JG. Prognostic Impact of Tumor-Infiltrating Lymphocytes and Neutrophils on Survival of Patients with Upfront Resection of Pancreatic Cancer. Cancers (Basel). 2019; 11.
37. Zhang J, Wang J, Qian Z, Han Y. CCR5 is Associated With Immune Cell Infiltration and Prognosis of Lung Cancer. J Thorac Oncol. 2019; 14: e102-e3.
38. Schmidt M, Bohm D, von Torne C, Steiner E, Puhl A, Pilch H, Lehr HA, Hengstler JG, Kolbl H, Gehrmann M. The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res. 2008; 68: 5405-13.
39. Iglesia MD, Vincent BG, Parker JS, Hoadley KA, Carey LA, Perou CM, Serody JS. Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. Clin Cancer Res. 2014; 20: 3818-29.
40. Germain C, Gnjatic S, Tamzalit F, Knockaert S, Remark R, Goc J, Lepelley A, Becht E, Katsahian S, Bizouard G, Validire P, Damotte D, Alifano M, et al. Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. Am J Respir Crit Care Med. 2014; 189: 832-44.
41. Fristedt R, Borg D, Hedner C, Berntsson J, Nodin B, Eberhard J, Micke P, Jirstrom K. Prognostic impact of tumour-associated B cells and plasma cells in oesophageal and gastric adenocarcinoma. J Gastrointest Oncol. 2016; 7: 848-59.
42. Berntsson J, Nodin B, Eberhard J, Micke P, Jirstrom K. Prognostic impact of tumour-infiltrating B cells and plasma cells in colorectal cancer. Int J Cancer. 2016; 139: 1129-39.
43. Spear S, Candido JB, McDermott JR, Ghirelli C, Maniati E, Beers SA, Balkwill FR, Kocher HM, Capasso M. Discrepancies in the Tumor Microenvironment of Spontaneous and Orthotopic Murine Models of

Pancreatic Cancer Uncover a New Immunostimulatory Phenotype for B Cells. Front Immunol. 2019; 10: 542.

Figures

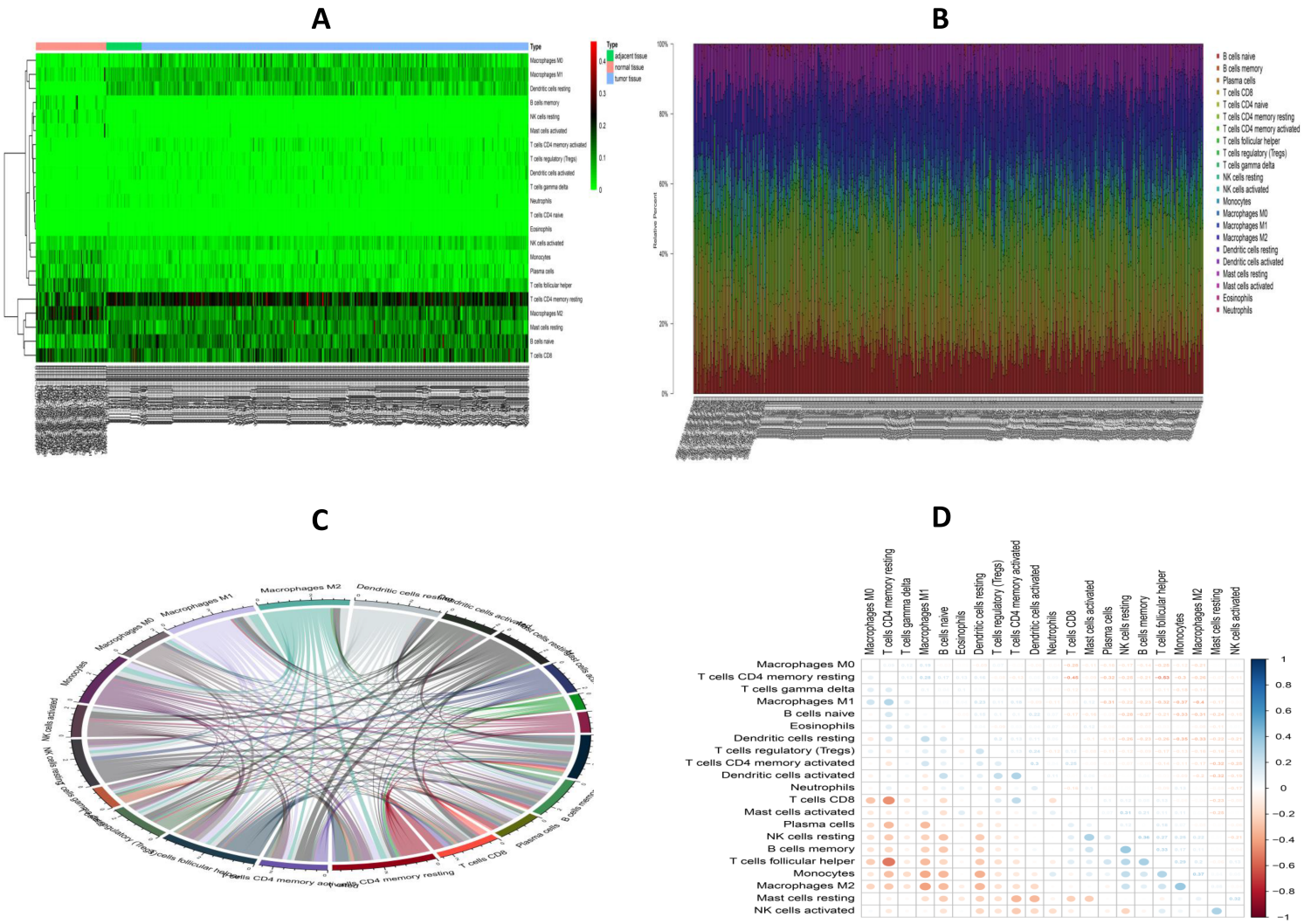

Figure 1

The performance of CIBERSOST for estimating TIICs composition in prostate cancer (PCa), adjacent tumor and control tissues. A. Heat map of 22 immune cell proportions. B. Relative proportions of 22 TIICs subpopulation. C and D. Correlation matrix of 22 immune cell densities in the TCGA cohort.

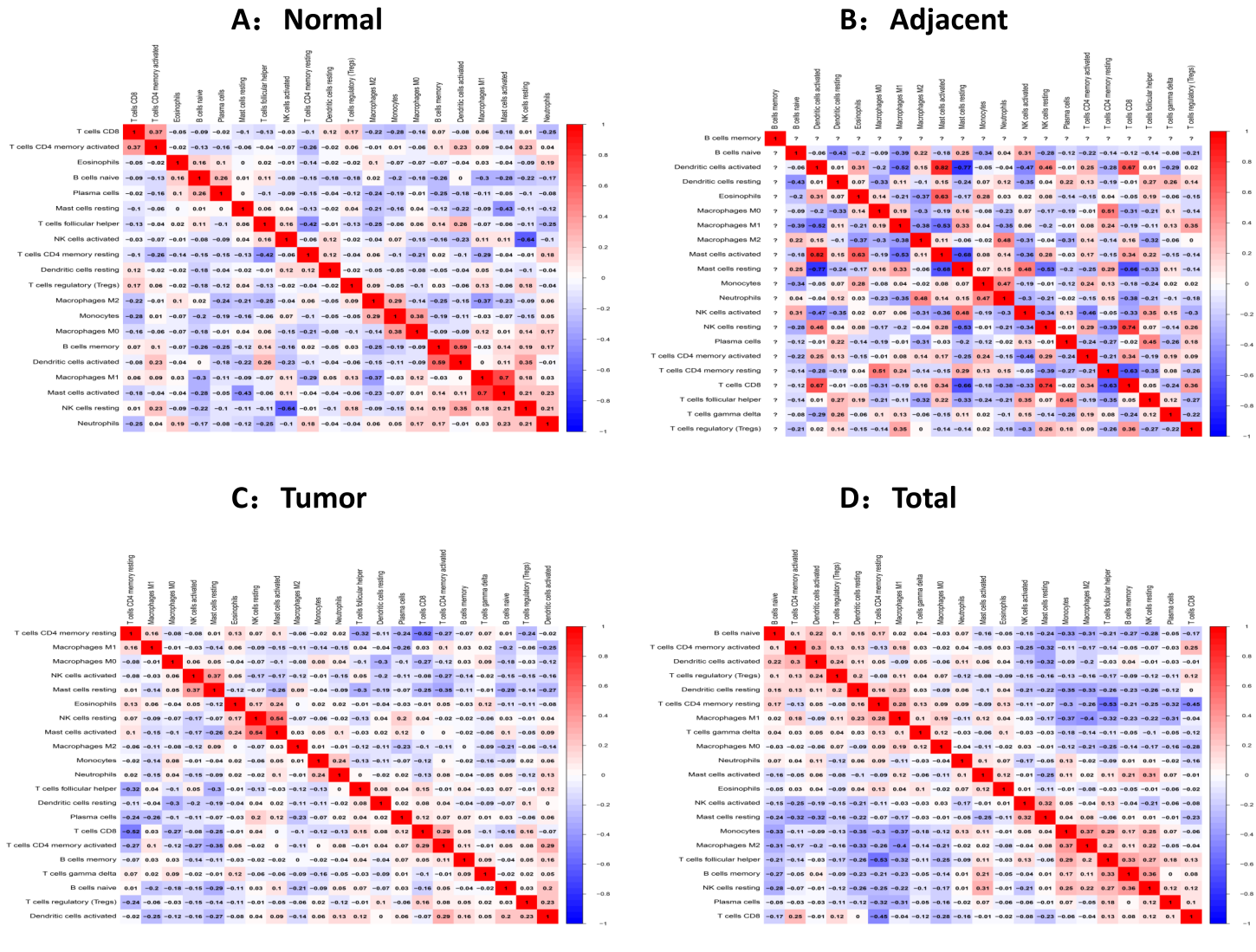

Figure 2

Correlation matrix of 22 immune proportions in the TCGA PCa cohort, including normal samples (A), adjacent tumor samples (B), tumor samples (C) and total samples (D).

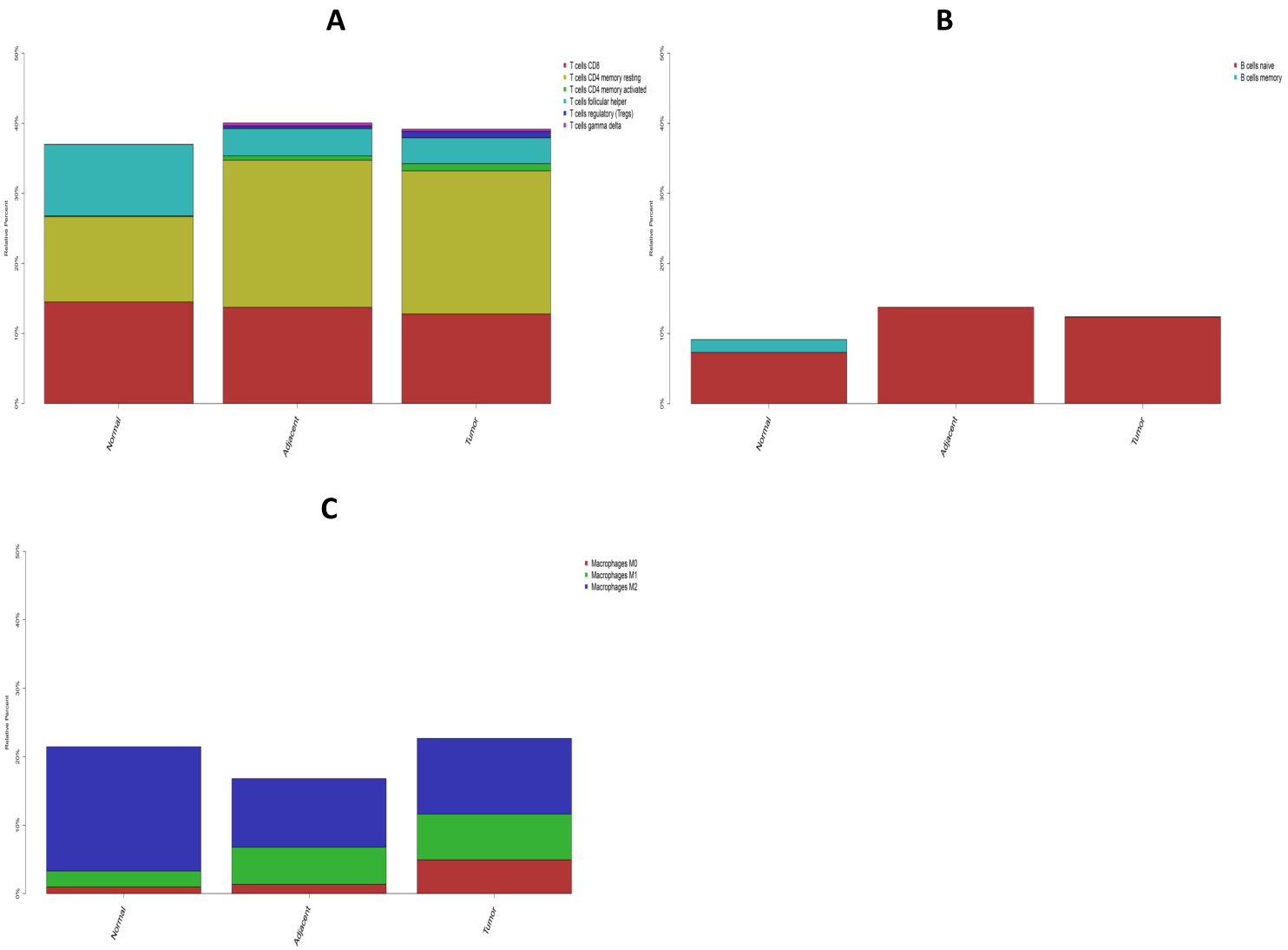

Figure 3

The stacked histogram exhibits the distribution of 22 immune cell infiltration among normal tissues, adjacent tumor tissues and tumor tissues, including total T cells (A), total B cells (B), and total macrophages (C).

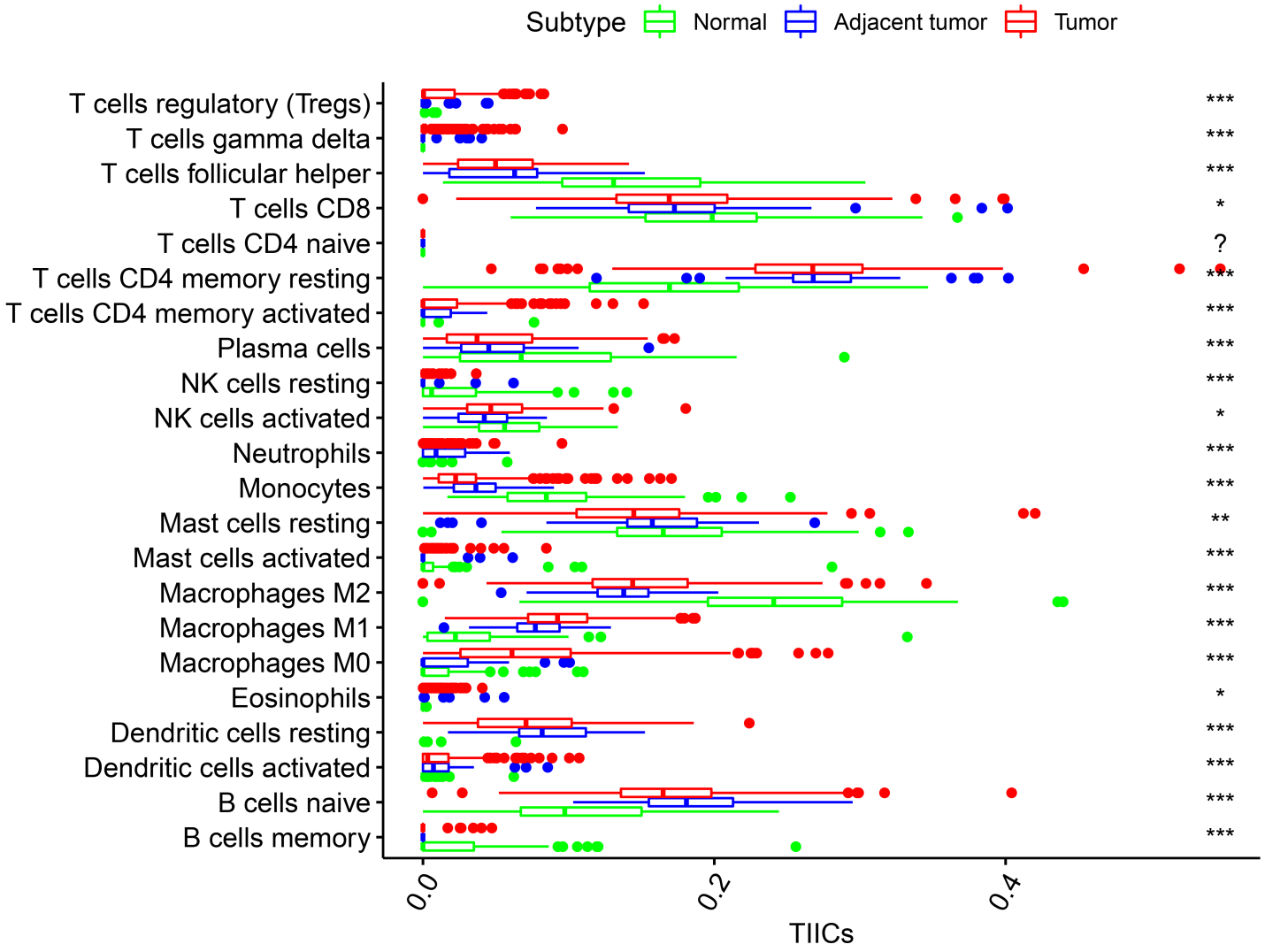

Figure 4

he Box plot shows the difference between CIBERSOFT immune cell fractions and normal tissues, adjacent tumor tissues and tumor tissues.

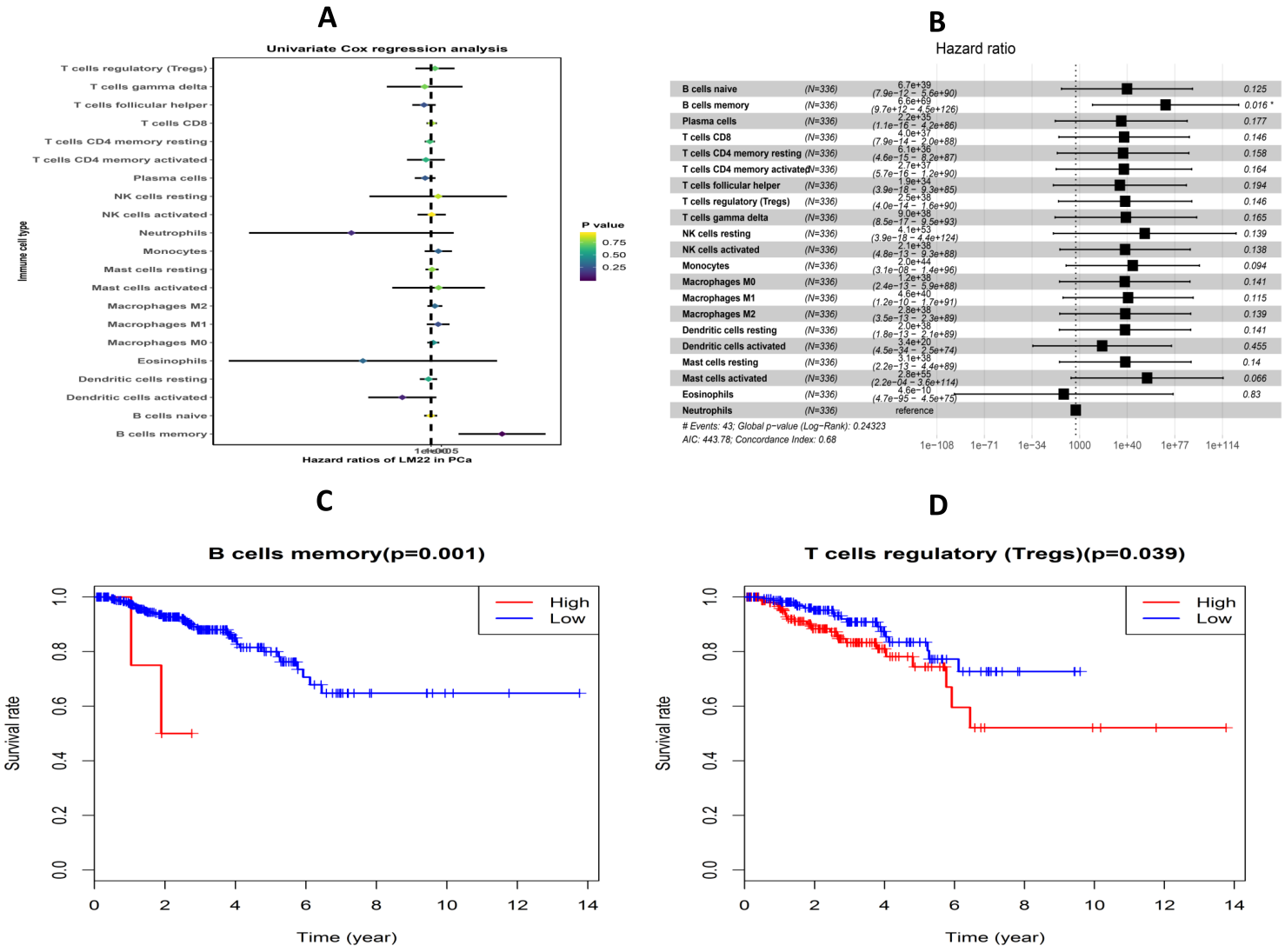

Figure 5

The identification of prognostic immune cells in TCGA PCa cohort, including the univariate Cox regression analysis (A) and multivariate Cox regression analysis (B). The K-M survival analysis of two immune cells, including B cells memory (A) and T cells regulatory (Tregs) (B).

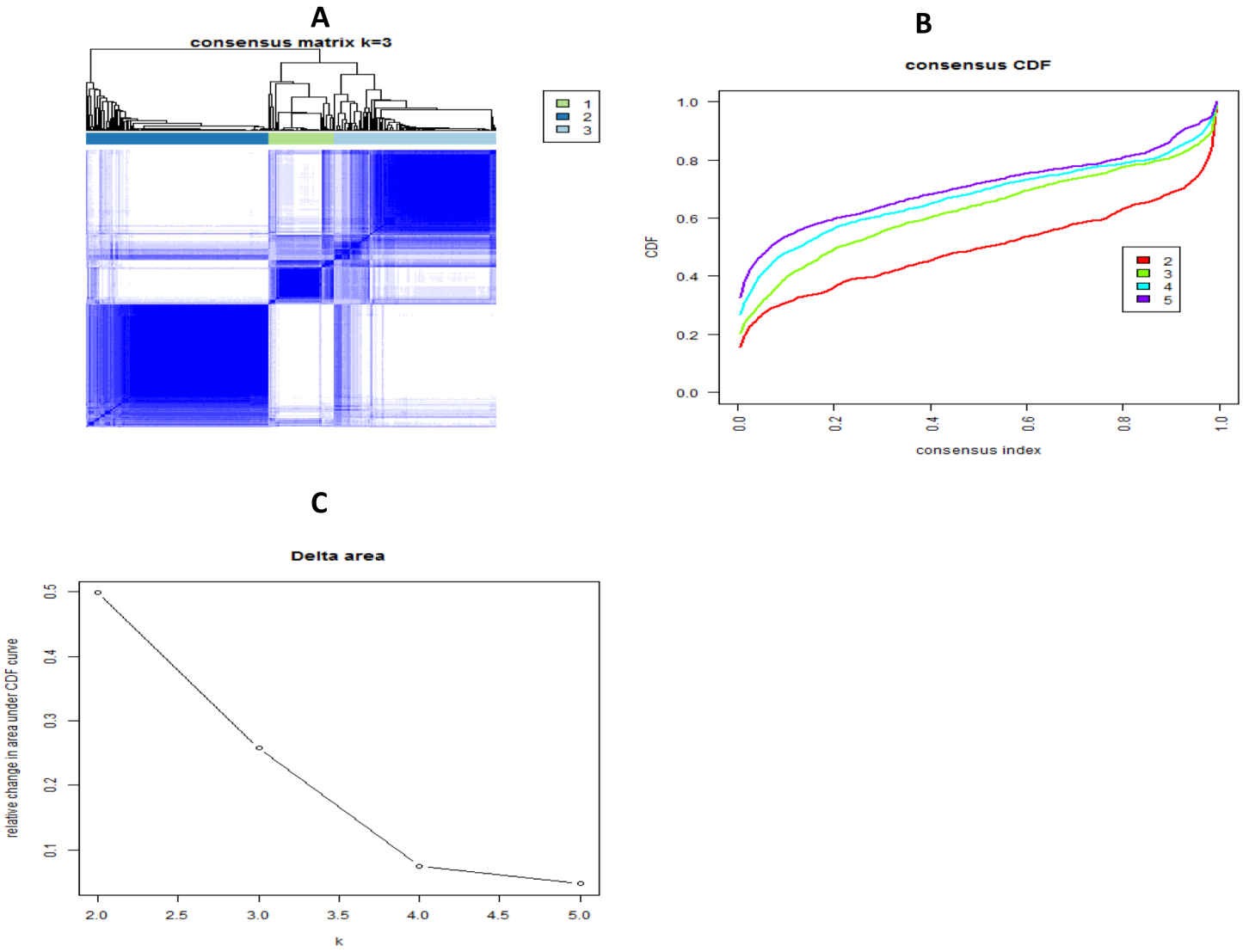

Figure 6

The cluster counts evaluated using Unsupervised clustering analysis. (A) The Consensus heatmap; (B) The relative change in area under the CDF curve of K= 2–5; (C) CDF curve of K=2–5.

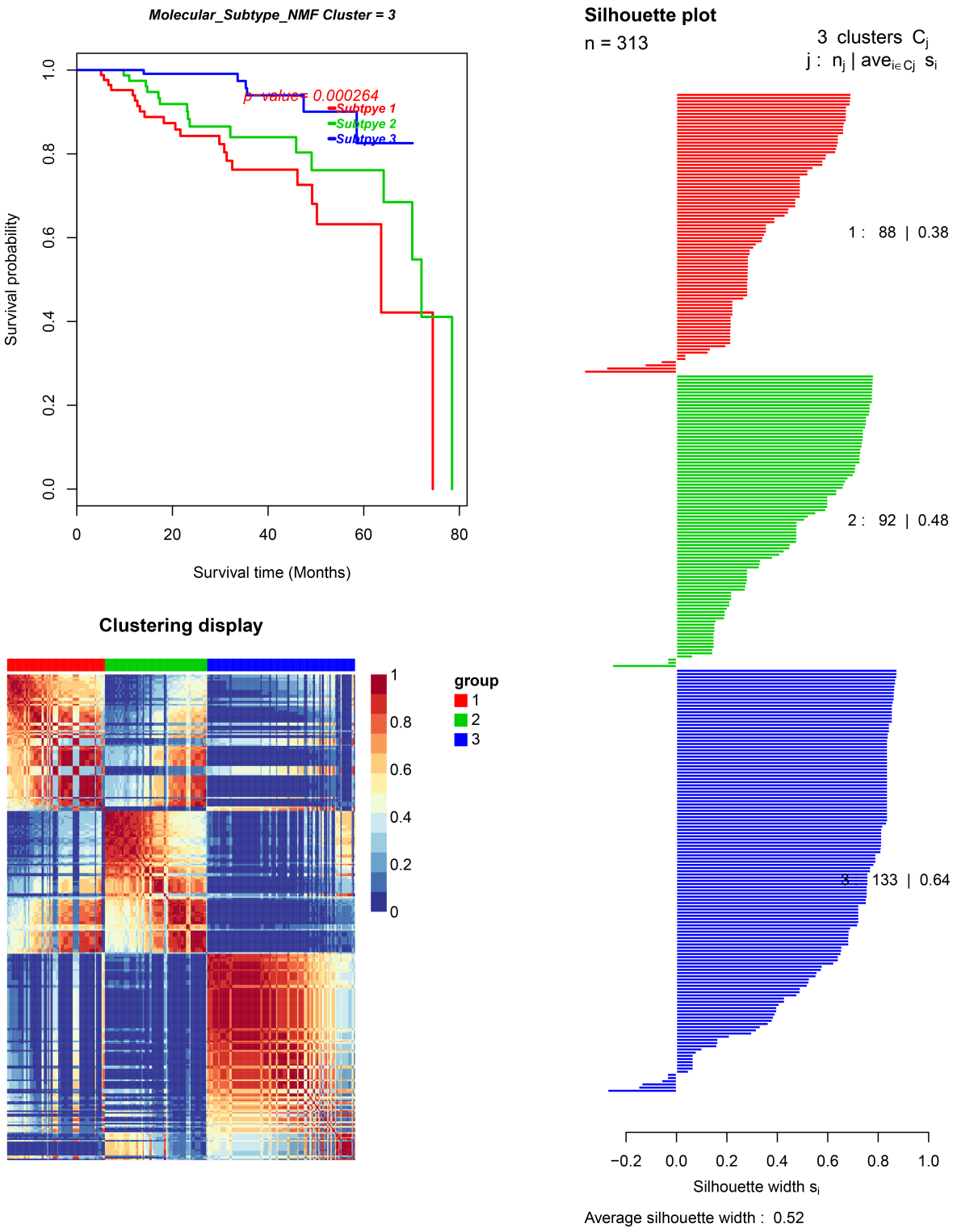

Figure 7

The identification of PCa subtypes using SNFCC+ algorithm. A. The log-rank test p-value for Kaplan–Meier survival analysis, B. The clustering heatmap visualizing the degree of the partitioning of the sample clusters, C. The average silhouette width representing the coherence of clusters.

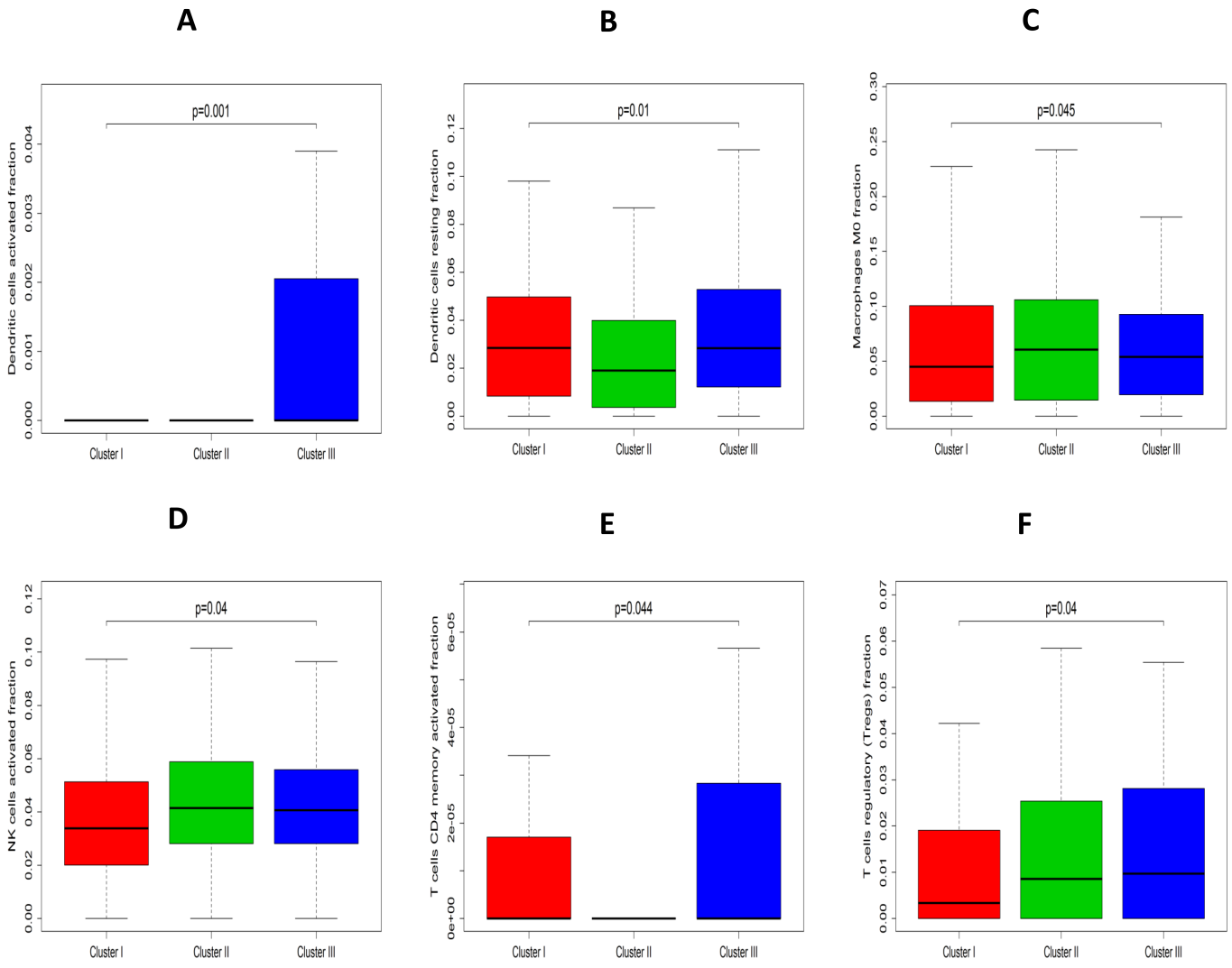

Figure 8

Box plot depicts the association between immune cell and three PCa clusters. A-F is for Dendritic cells activated, Dendritic cells resting, Macrophages M0, NK cells activated, T cells CD4 memory activated, and T cells regulatory (Tregs).

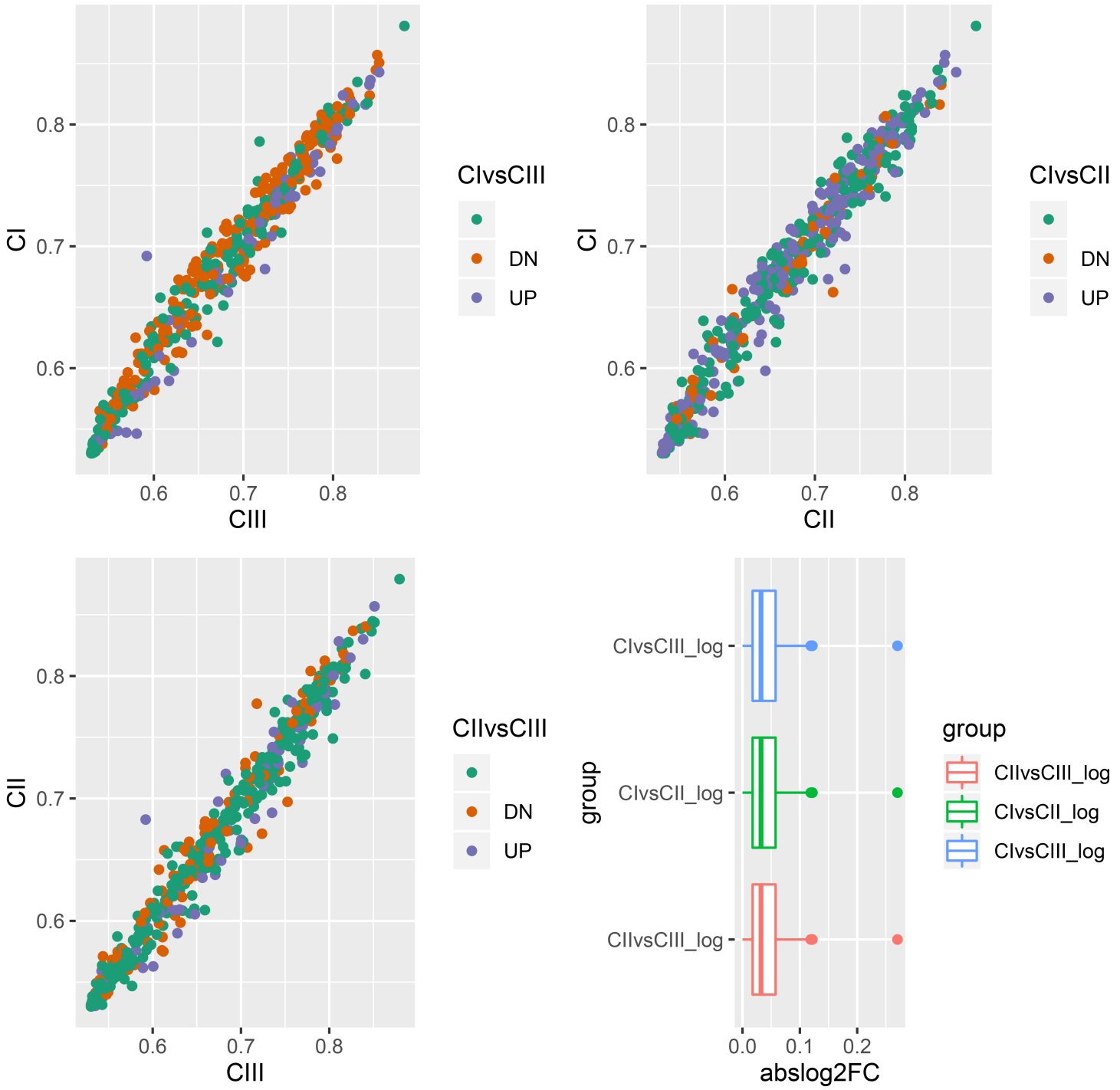

Figure 9

Differential gene expression analysis for immune related genes. Each dot-plot shows all immune related genes, those with upregulated expression (red), and those with downregulated expression (blue). Bottom right panel: boxplot showing the distribution of absolute fold changes in the three comparisons made.

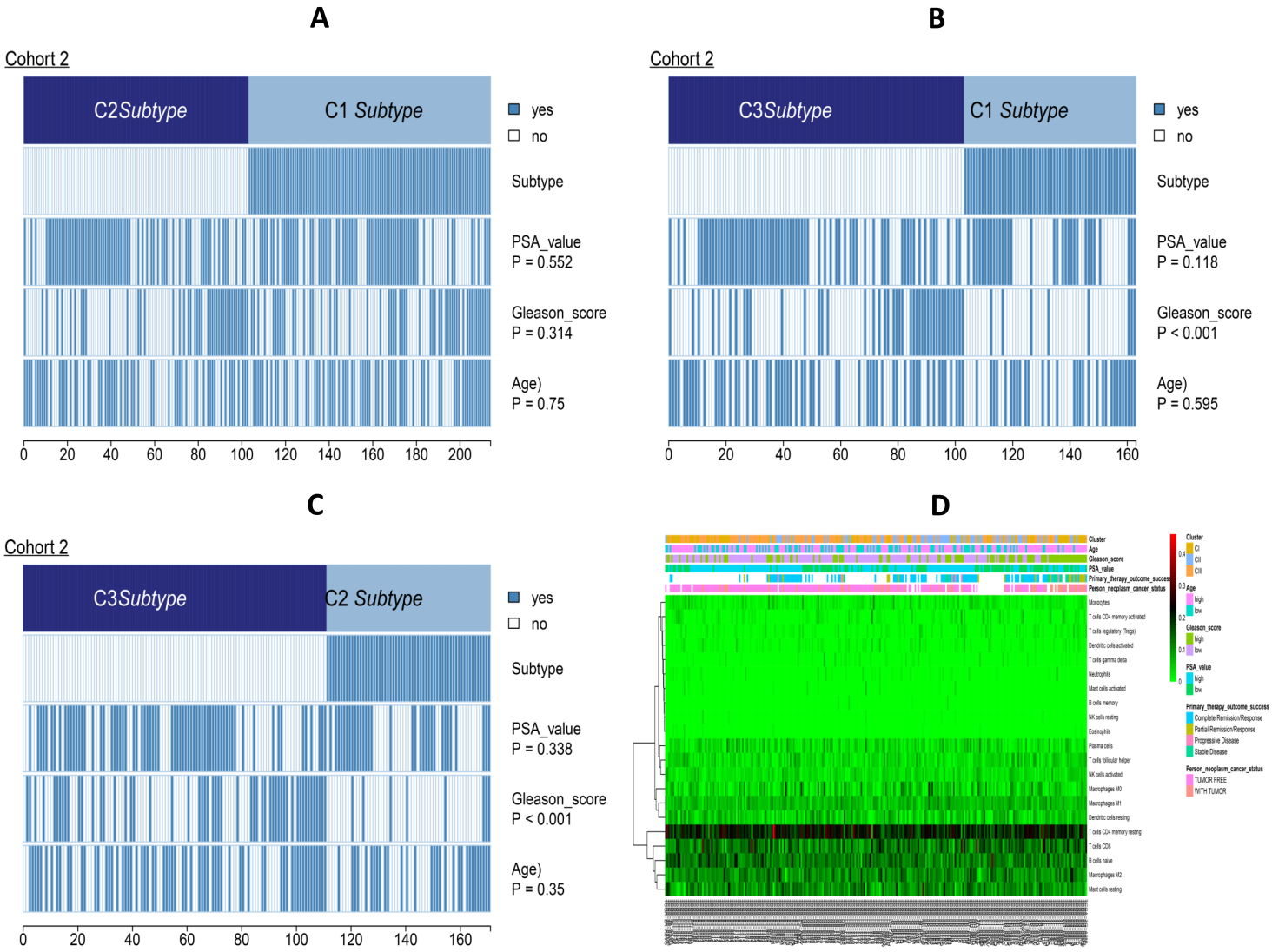

Figure 10

Comparing PSA value, Gleason score, and Age between PCa subsets. Statistical significance was performed by Chi-square test. A,B,C. is for CI vs CII, CI vs CIII, CII vs CIII, respectively. The heatmap illustrates the association of different clinical characters with PCa subsets (D).

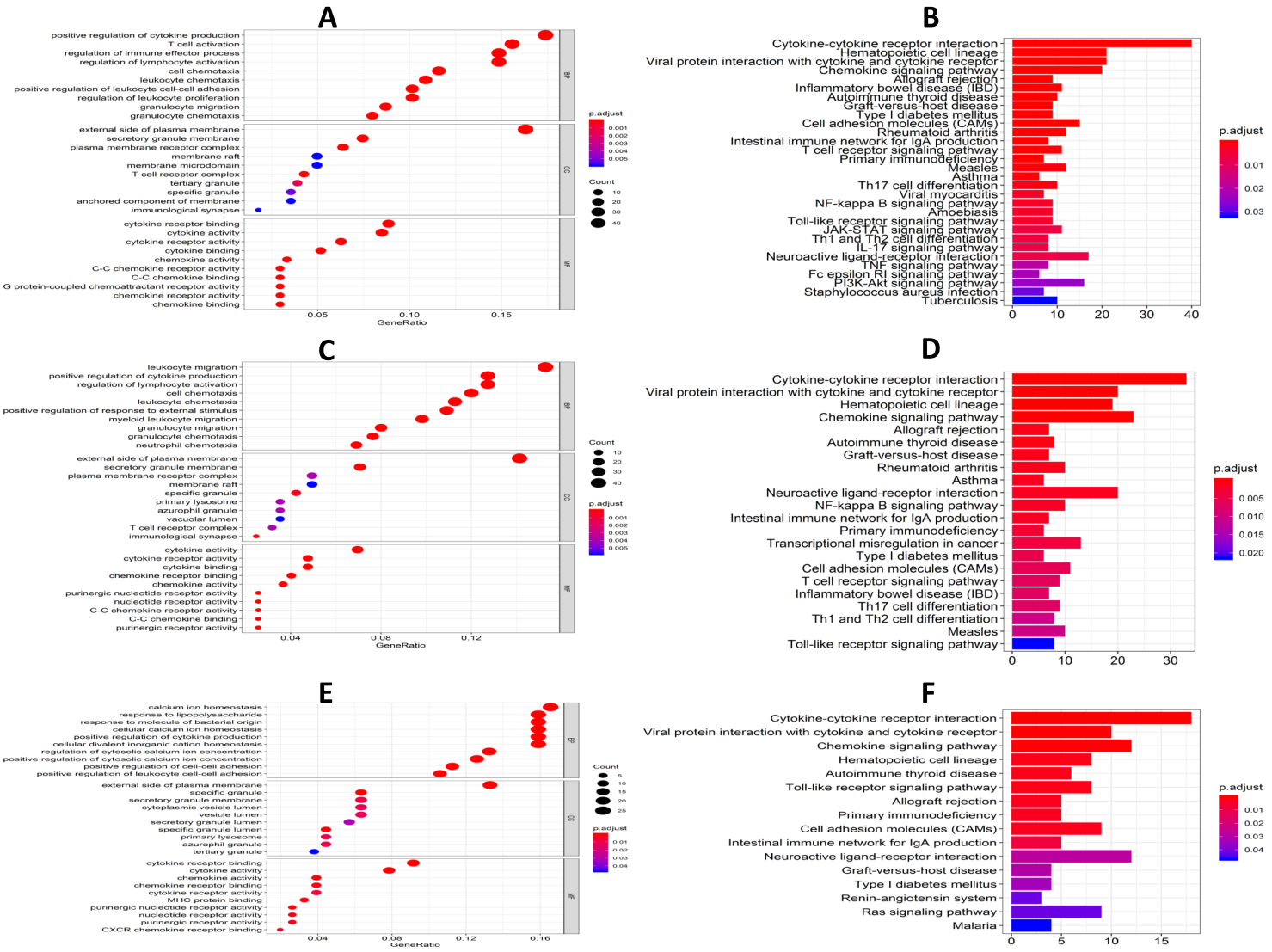

Figure 11

The GO and KEGG analysis for three PCa clusters. A,B is for cluster I vs Cluster II, C,D is for cluster I vs cluster III, and E,F is for cluster II vs Cluster III.

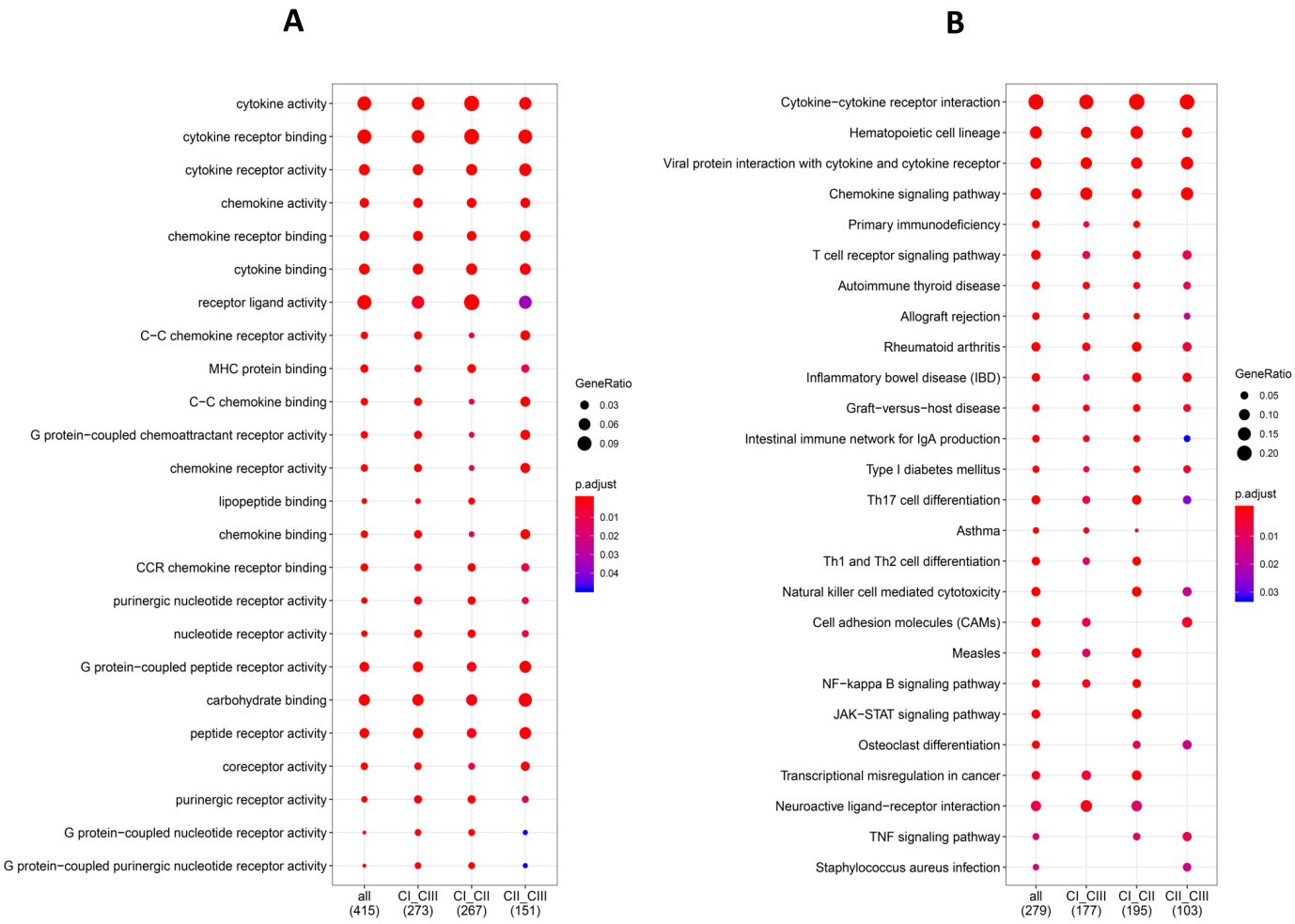

Figure 12

The comparision of functional enriched analysis among three PCa subsets, including GO (A) and KEGG (B) analysis.

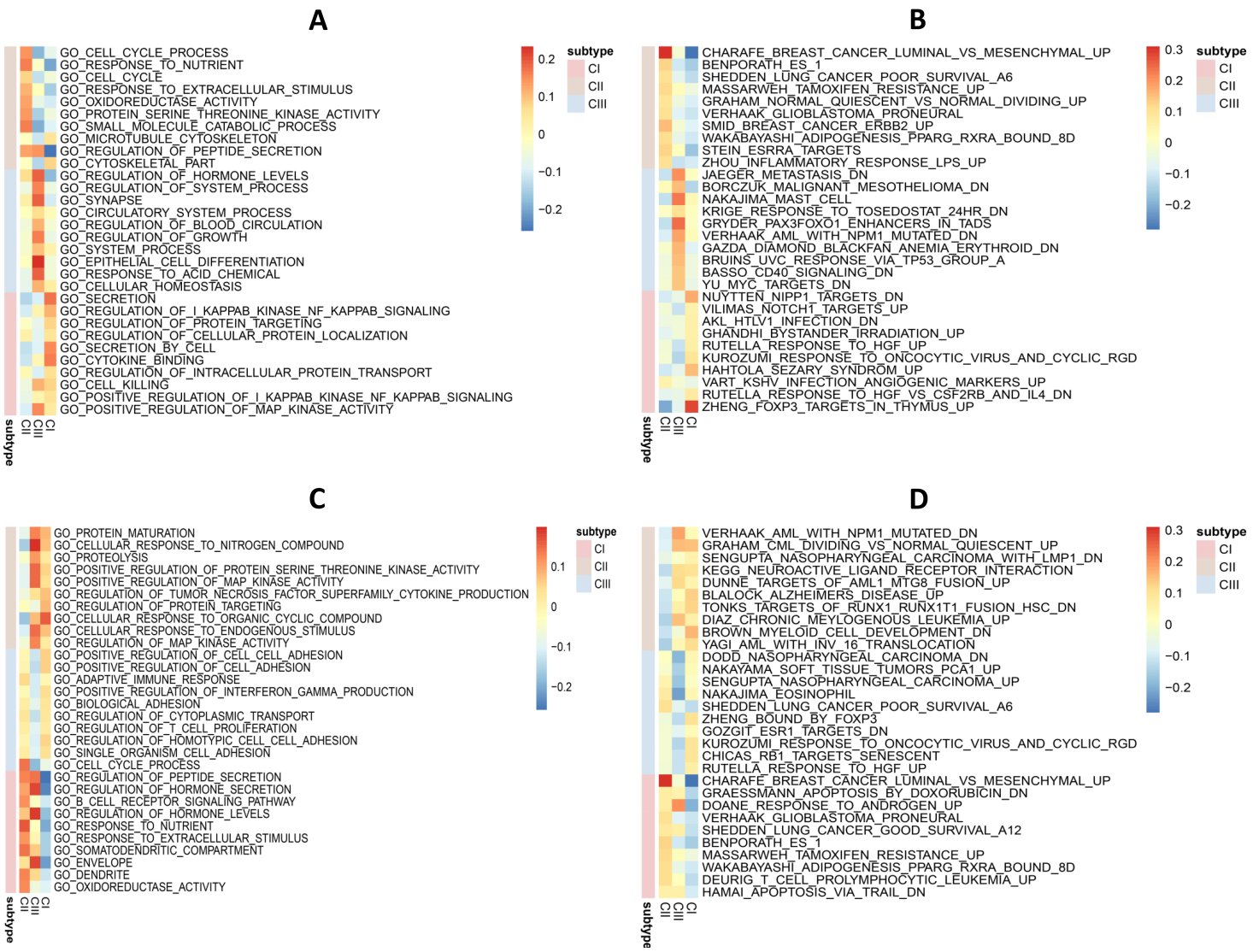

Figure 13

GSEA analysis reveals distinct enriched gene sets between three PCa subtypes, including up-regualted GO (A) and KEGG (B) categories, down-regualted GO (C) and KEGG (D) terms.

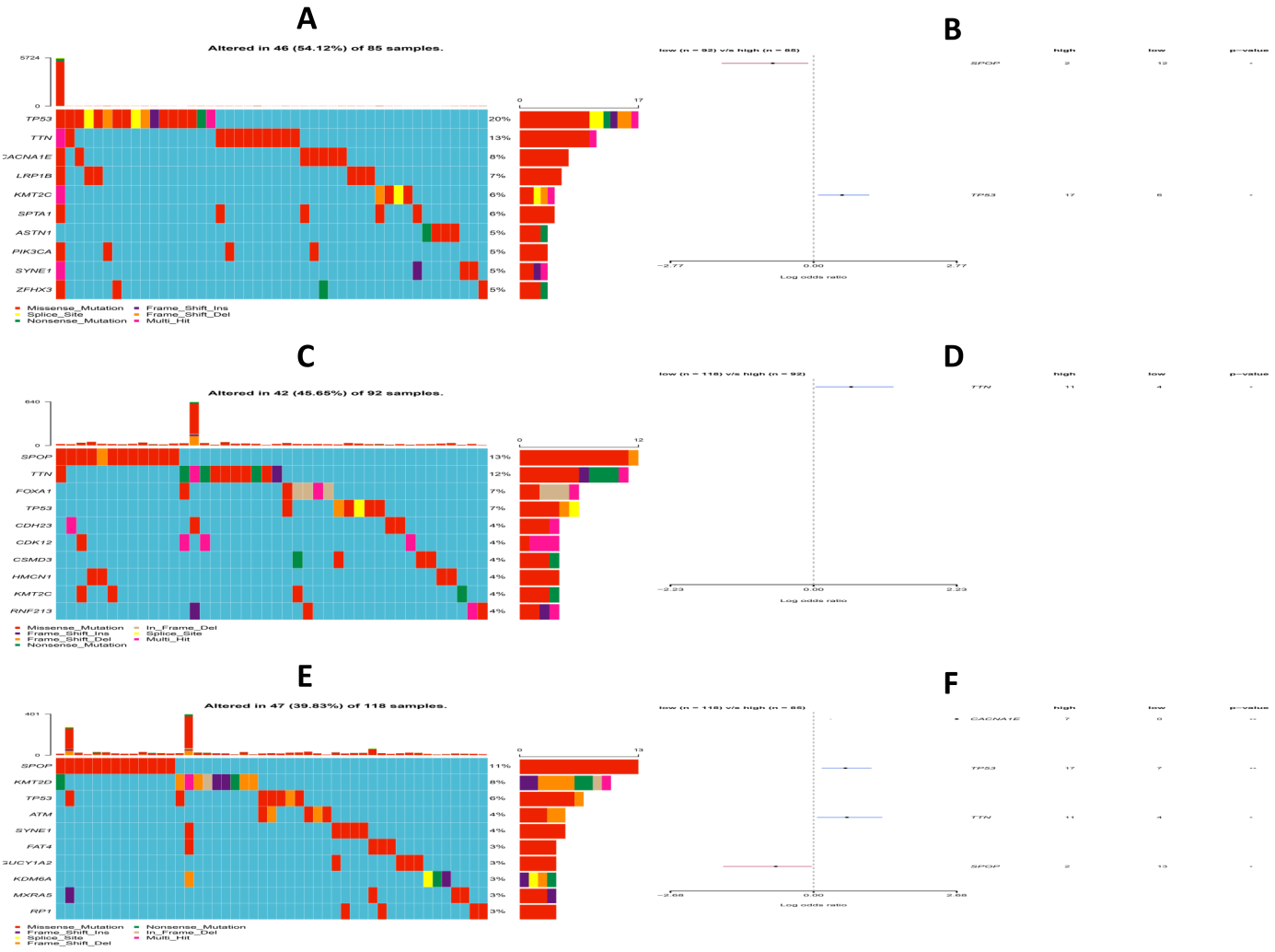

Figure 14

Mutation analysis of three PCa subsets. A,C,E. Gene mutation profiles of the following highly mutated genes among the three subtypes. B,D,F. The forest plots show the comparison results of gene mutations among CI, CII and CIII (*P, 0.1, **P, 0.05, ns: not significant).

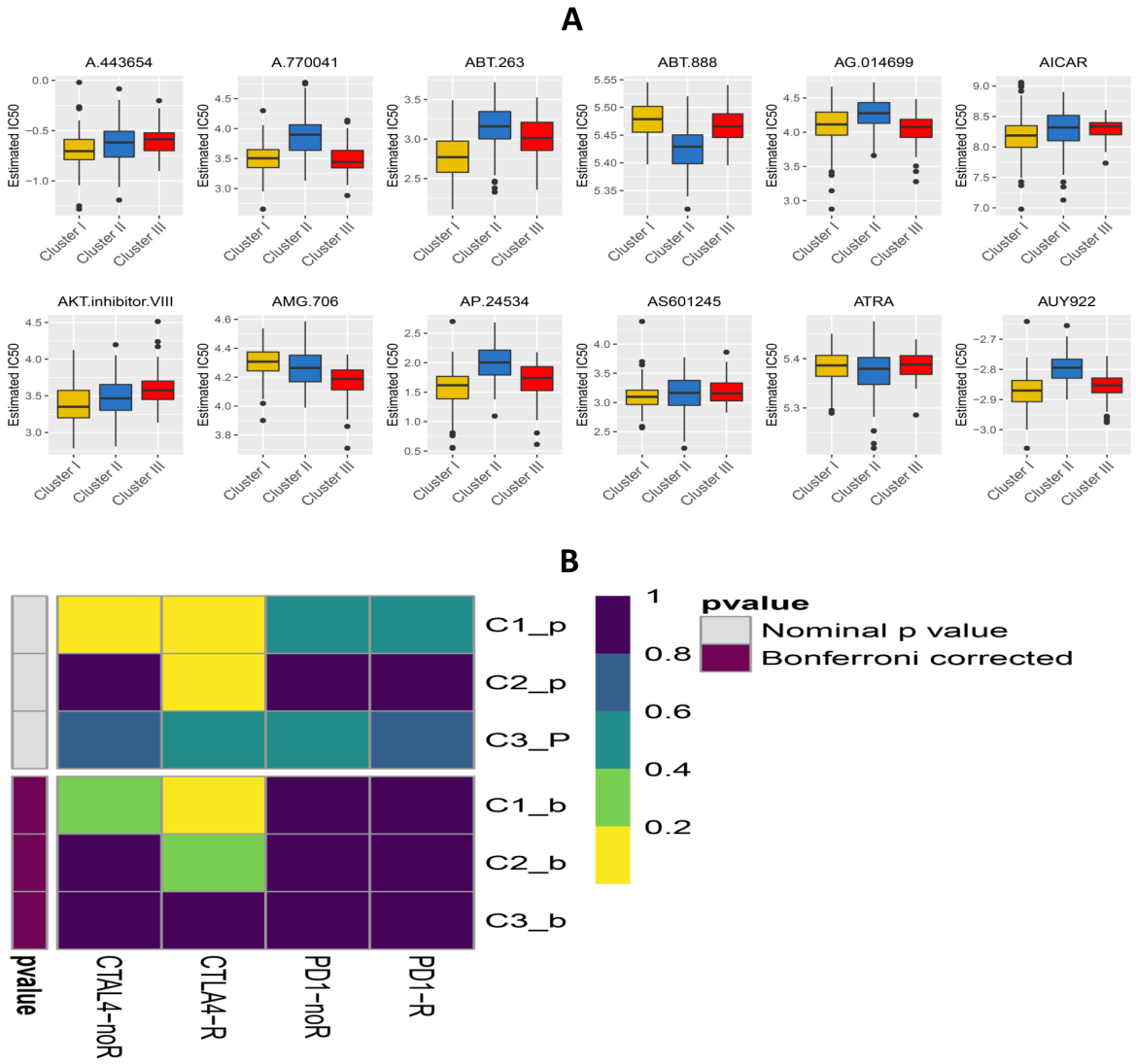

Figure 15

Differential putative chemotherapeutic and immunotherapeutic response. The box plots of the estimated IC 50 for chemotherapeutic drugs are shown in (A) for CI, CII and CIII PCa subsets. (C) Submap analysis manifested that Cluster I could be more sensitive to the immunotherapy (Bonferroni-corrected P < 0.05).