Article

Comprehensive analysis of CPA4 as a poor prognostic biomarker correlated with immune cells infiltration in bladder cancer

Chengcheng Wei [1, †], Yuancheng Zhou [1, †], Qi Xiong [2, †], Ming Xiong[1,] *and Zhaohui Chen [1,] *

1   Department of Urology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; chengchengwei@hust.edu.cn (C.W); 1012660651@qq.com (Y.Z)
2   Chongqing Medical University, Chongqing, China; xq827928@163.com (Q.X)
*   Correspondence: Xiong_Ming@hust.edu.cn(M.X); zhaohuichen@hust.edu.cn(Z.C)
† These authors contributed equally to this work.

Simple Summary: Overexpression of Carboxypeptidase A4 (CPA4) has been observed in plenty types of cancer and elucidated to promote tumor growth and invasion; however, its role in bladder urothelial carcinoma (BLCA) is still unclear. Therefore, we aimed to show the prognostic role of CPA4 and its relationship with immune infiltrates in BC. we confirmed that overexpression of CPA4 is associated with shorter overall survival, disease specific survival, progress free interval and higher dead event. Moreover, we unraveled that several infiltrating immune cells (Th1cell, Th2 cell, T cell exhaustion and Tumor-associated macrophage) were correlated with the expression of CPA4 in bladder cancer by using TIMER2 and GEPIA2. In conclusion, CPA4 may be a novel and great prognostic biomarker based on the bioinformation analysis in BLCA.

Abstract: Carboxypeptidase A4 (CPA4) has shown the potential possibility as a biomarker in the early diagnosis for certain cancers. However, no previous research has linked CPA4 to therapeutic or prognostic significance in bladder cancer. Using data from The Cancer Genome Atlas (TCGA) database, we set out to determine the full extent of the link between CPA4 and BLCA. We further analyzed the interacting proteins of CPA4 and infiltrated immune cells via TIMER2，STRING and GEPIA2 databases. The expression of CPA4 in tumor and normal tissues was compared using the TCGA+GETx database. The connection between CPA44 expression and clinicopathologic characteristics and overall survival (OS) was investigated using multivariate methods and Kaplan-Meier survival curves. The potential functions and pathways were investigated via gene set enrichment analysis. Furthermore, we analyze the associations between CPA4 expression and infiltrated immune cells with their respective gene marker sets using the ssGSEA, TIMER2, and GEPIA2 databases. Compared to matching normal tissues, human CPA4 was found to be substantially expressed. We confirmed that overexpression of CPA4 is linked with shorter OS, DSF, PFI, and increased diagnostic potential using Kaplan-Meier and ROC analysis. The expression of CPA4 is related to T-bet, IL12RB2, CTLA4, and LAG3, among which T-bet and IL12RB2 are Th1 marker genes, while CTLA4 and LAG3 are related to T cell exhaustion, which may be used to guide the application of checkpoint blockade and the adoption of T cell transfer therapy.

Keywords: CPA4, Bladder Urothelial Carcinoma, Immune cells, T cell exhaustion, checkpoint

1. Introduction

Bladder Urothelial Carcinoma (BLCA) is the 8th most prevalent cancer worldwide, with 549393 new cases reported worldwide in 2018[1]. And in the USA alone, there are estimated more than 80000 new cases and 17000 deaths each year[2]. This disease is particularly heterogeneous.[3] They are classified as high-grade and low-grade diseases based on standardized histomorphologic features described by the World Health Organization. The invasion depth of bladder wall determines the tumor staging. Approximately 80% of BC patients present non-muscle-invasive bladder cancer (NMIBC) at the time of diagnosis, while the remainder presented muscle-invasive bladder cancer (MIBC) or even distant metastases.[4] NMIBCs do not normally pose a threat to patients' survival and have a much better prognosis due to effective therapeutic options.[5] However, they almost always relapse, and patients need to repeat intravesical treatments, endoscopic evaluations and biopsies, which may take an extended period of time, resulting expensive surgical and surveillance management[6-8]. MIBCs, on the other hand, are clinically aggressive and can progress rapidly to lymph nodes, brain, lungs, liver, bone metastases, which are often fatal[3]. However, over the past three decades, clinical management and 5-year survival rates had seen little substantial advances[9]. Therefore, it is significant to identify novel biomarkers and molecular targets for advancing the prognosis of bladder cancer (BC).

Carboxypeptidase A4 (CPA4) is a member of the zinc-containing metallocarboxypeptidase family[10], which could specifically catalyze the peptide bonds released from carboxy-terminal amino acids[11,12]. CPA4 was first discovered when screening for the upregulated mRNA during cancer cell differentiation induced by sodium butyrate[13]. From the cellular and biochemical characteristics, CPA4 is secreted from cells in the form of soluble proenzyme (pro-CPA4), which might play a role in creating tumor microenvironment[10]. Previous studies have demonstrated that CPA4 is closely associated with the aggressiveness, growth and differentiation in cancer cells[14,15]. But the underlying mechanism of CPA4 in BLCA remains unclear.

Recently, CPA4 has shown the potential possibility as a biomarker in the early diagnosis for certain cancers. Sun et al. have reported that the higher expression level of CPA4 in pancreatic cancer tissues and serum is related to poor prognosis and higher aggressiveness [13]. Previously studied showed that upregulated mRNA level of CPA4 in androgen-independent prostate cancer cells is associated with the Histone Hyperacetylation signaling pathway[16].In liver cancer and lung cancer, studies have also shown that the higher expression of CPA4 was closely associated with the early diagnosis and poor prognosis[13,17]. Despite the potential significance of CPA4 expression in plenty types of cancer, there are no previous studies have ever shown the expression level of CPA4 in bladder cancer, and its potential therapeutic and prognostic values. Also, the correlation with immune infiltrates of CPA4 in bladder cancer remains to be investigated. Shao et al. demonstrated that CPA4 overexpression promotes the progression of aggressive clinical stage in pancreatic cancer and the downregulation of CPA4 inhibits non-small-cell lung cancer growth[15,18]. So, we hypothesized that the level of CPA4 is associated with the prognosis and immune cells infiltration in BC.

To test this hypothesis, our study evaluated the role of CPA4 on tumorigenesis and clinical significance based on The Cancer Genome Atlas (TCGA). We have compared the different expression level of BC in age, gender, pathologic T, N, M stage, pathology, subtype and OS. In this study, we found that CPA4 is upregulated in BC. Significantly, the risk factors of CPA4 upregulation are correlated with poor prognosis. And the correlation with immune infiltrates of CPA4 for BC is also evaluated. Eventually, we link the high CPA4 level and poor prognosis in BC.

2. Materials and Methods

2.1 Data source

The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) provides 33 types of cancer clinic and pathological information for scholars and researchers for free[19]. The expression profiles of CPA4 and clinical information of TCGA cancer data were downloaded from the UCSC Xena (https://xenabrowser.net/datapages/) database. TCGA database is publicly open-access and available where ethical approval and informed consent of the patients were not necessary.

2.2 CPA4 Methylation Level Analysis

UALCAN (http://ualcan.path. uab.edu/) is a comprehensive, user-friendly, and interactive web resource for analyzing cancer OMICS data providing graphs and plots depicting expression profile and patient survival information for protein-coding, miRNA-coding and lincRNA-coding genes[20]. The UALCAN online tool was utilized to analyze the CPA4 methylation level in BLCA (TCGA data).

2.3 Analysis of Differentially Expressed Genes (DEGs)

Through the limma Package by R, patients with different CPA4 expression profiles in High and Low Expression Groups (HTSeq-TPM) were compared using unpaired Student's t-test to identify the DEGs[21]. A $|\log 2\text{Fold Change}| > 2$ and adjusted $P < 0.05$ were considered the threshold for the DEGs. Gene Ontology (GO) Enrichment Analysis. Metascape (https://metascape.org) is a means for gene annotation and pathway analysis[22]. In this study, Metascape was utilized to analyze the enrichment of CPA4 related DEGs by process and pathway. Terms with a $P < 0.01$, a minimum count of 3, and an enrichment factor of >1.5 were regarded as significant.

2.4 Gene set enrichment analysis (GSEA)

GSEA used statistical method in order to seek whether gene exhibits are statistically significant concordant between two biological states[23]. We used R package Cluster Profiler to evaluate the excessive function and pathway differences between the groups with different expression of CPA4 expression[24]. Each analysis of process repeated 1000 times. Adjusted P-value < 0.05, and false discovery rate (FDR) < 0.25 was considered statistically significant enrichment

2.5 Comprehensive analysis of Protein-protein interaction

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) website (https://string-db.org/) is a database of knew and predicted protein-protein interactions which host a collection of integrated and consolidated protein-protein interaction data including direct (physical) and indirect (functional) associations[25]. By importing the CPA4 into the online tool STRING, protein–protein interaction (PPI) network information was complied. The confidence scores > 0.4 was considered median significant.

2.6 Analysis of The Tumor Immune Estimation Resource (TIMER2)

The Tumor Immune Estimation Resource (TIMER2) is a comprehensive resource including 32 cancer types and incorporate 10,897 samples from TCGA database, for systematically analysis of immune infiltrates across diverse cancer types (http://cistrome.org/TIMER/)[26]. TIMER2 database evaluated the correlation of expression of CPA4 in BLCA patients with the six types of infiltrating immune cells (B cells, Dendritic cells, CD4+T cells, CD8+T cells, Macrophages, and Neutrophils) and displayed the relationship between the expression of the CPA4 gene and the tumor purity.

2.7 Immune infiltration analysis by ssGSEA

Single sample GSEA (ssGSEA) was performed to analyze the state of immune infiltration of BLCA from R package GSVA (version3.6) (http://www.bioconductor.org/ packages/release/bioc/html/GSVA.Html), and we quantified infiltration levels of 24 immune cell types from gene expression profile in the literature[27]. In order to discover the

correlation between CPA4 and the infiltration levels of 24 immune cells, P values were established by the Spearman and Wilcoxon rank-sum test.

2.8 Analysis of The Gene Expression Profiling Interactive Analysis 2

The Gene Expression Profiling Interactive Analysis2 (GEPIA2) (http://gepia.cancer-pku.cn/index.html) is an updated database which for analyzing the RNA sequencing expression data of 9,736 tumors and 8,587 normal samples from the TCGA and the GTEx projects which include 60498 genes and 198619 isoforms[28]. GEPIA2 database investigated the expression level of CPA4 with various immune cells' markers. TIMER2 was used to identify the gene which of significant correlation with CPA4 expression in the GEPIA2 web.

2.9 Statistical Analysis

The expression of CPA4 of non-paired and paired samples was analyzed by the Wilcoxon rank-sum test and Wilcoxon signed-rank test respectively. By using pROC package to generate the ROC curve to evaluate the CPA4 expression with diagnostic performance. The relations between the CPA expression and the clinical features were analyzed by the Kruskal–Wallis test, Chi-Squared test and Wilcoxon signed rank test. By using Kaplan–Meier to generate the survival curves via long-rank test. By using Cox proportional hazard modeling to draw the university and multivariate estimating the risk of death $P < 0.05$ was considered statistically significant: $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; R software processed all kinds of statistical analysis (Version 4.0.2)

3. Results

3.1 Characteristics of BLCA patients

Totally, 414 BLCA tumor tissues and 19 normal tissues information have been collected from the TCGA database including RNA-seq and relative clinical prognostic information in 414 patients. We grouped the BLCA patients into two set including low (n=207) and high expression(n=207) of CPA4. Clinical information of BLCA patients including age, race, gender, pathologic stage, pathologic stage (T, N or M), Pathologic stage, primary therapy outcome, histologic grade, radiation therapy, subtype, smoker, lymphovascular invasion, OS event (**Table 1**)

Table 1. clinical characteristics of two sets of patients in different expression of CPA4 from TCGA dataset.

Characteristic	Low expression of CPA4	High expression of CPA4	p
n	207	207	
Age, n (%)			0.921
<=70	116 (28%)	118 (28.5%)	
>70	91 (22%)	89 (21.5%)	
Race, n (%)			0.003
Asian	32 (8.1%)	12 (3%)	
Black or African American	8 (2%)	15 (3.8%)	
White	159 (40.1%)	171 (43.1%)	
Gender, n (%)			0.372
Female	50 (12.1%)	59 (14.3%)	
Male	157 (37.9%)	148 (35.7%)	
T stage, n (%)			0.004
T1	4 (1.1%)	1 (0.3%)	
T2	73 (19.2%)	46 (12.1%)	
T3	89 (23.4%)	107 (28.2%)	
T4	23 (6.1%)	37 (9.7%)	
N stage, n (%)			0.494

N0	120 (32.4%)	119 (32.2%)	
N1	18 (4.9%)	28 (7.6%)	
N2	39 (10.5%)	38 (10.3%)	
N3	3 (0.8%)	5 (1.4%)	
M stage, n (%)			0.810
M0	109 (51.2%)	93 (43.7%)	
M1	5 (2.3%)	6 (2.8%)	
Pathologic stage, n (%)			0.014
Stage I	4 (1%)	0 (0%)	
Stage II	76 (18.4%)	54 (13.1%)	
Stage III	63 (15.3%)	79 (19.2%)	
Stage IV	63 (15.3%)	73 (17.7%)	
Radiation therapy, n (%)			0.369
No	181 (46.6%)	186 (47.9%)	
Yes	13 (3.4%)	8 (2.1%)	
Primary therapy outcome, n (%)			< 0.001
PD	18 (5%)	52 (14.6%)	
SD	14 (3.9%)	17 (4.8%)	
PR	12 (3.4%)	10 (2.8%)	
CR	136 (38.1%)	98 (27.5%)	
Histologic grade, n (%)			< 0.001
High Grade	186 (45.3%)	204 (49.6%)	
Low Grade	19 (4.6%)	2 (0.5%)	
Lymphovascular invasion, n (%)			0.666
No	62 (21.9%)	68 (24%)	
Yes	78 (27.6%)	75 (26.5%)	
Subtype, n (%)			0.003
Non-Papillary	124 (30.3%)	151 (36.9%)	
Papillary	82 (20%)	52 (12.7%)	
OS event, n (%)			< 0.001
Alive	139 (33.6%)	92 (22.2%)	
Dead	68 (16.4%)	115 (27.8%)	
Age, meidan (IQR)	69 (60, 76)	68 (61, 76)	0.990

3.2 Tumor tissues express higher CPA4 than normal tissue

The expression of CPA4 in pan-cancer was analyzed between tumor and normal tissues. From TCGA+GETx database, the expression level of CPA4 in non-matched (p=1.6e-5) was significantly higher than the normal (**Fig 1**). The analysis of correlation CPA4 expression of BLCA patients and relative clinical information which shows that higher DLEU1 expression level correlated with OS event, subtype of papillary. There was no any statistically significant difference between the expression levels of CPA4 in BLCA and age, gender, pathologic T, N, M stage and pathologic stage.
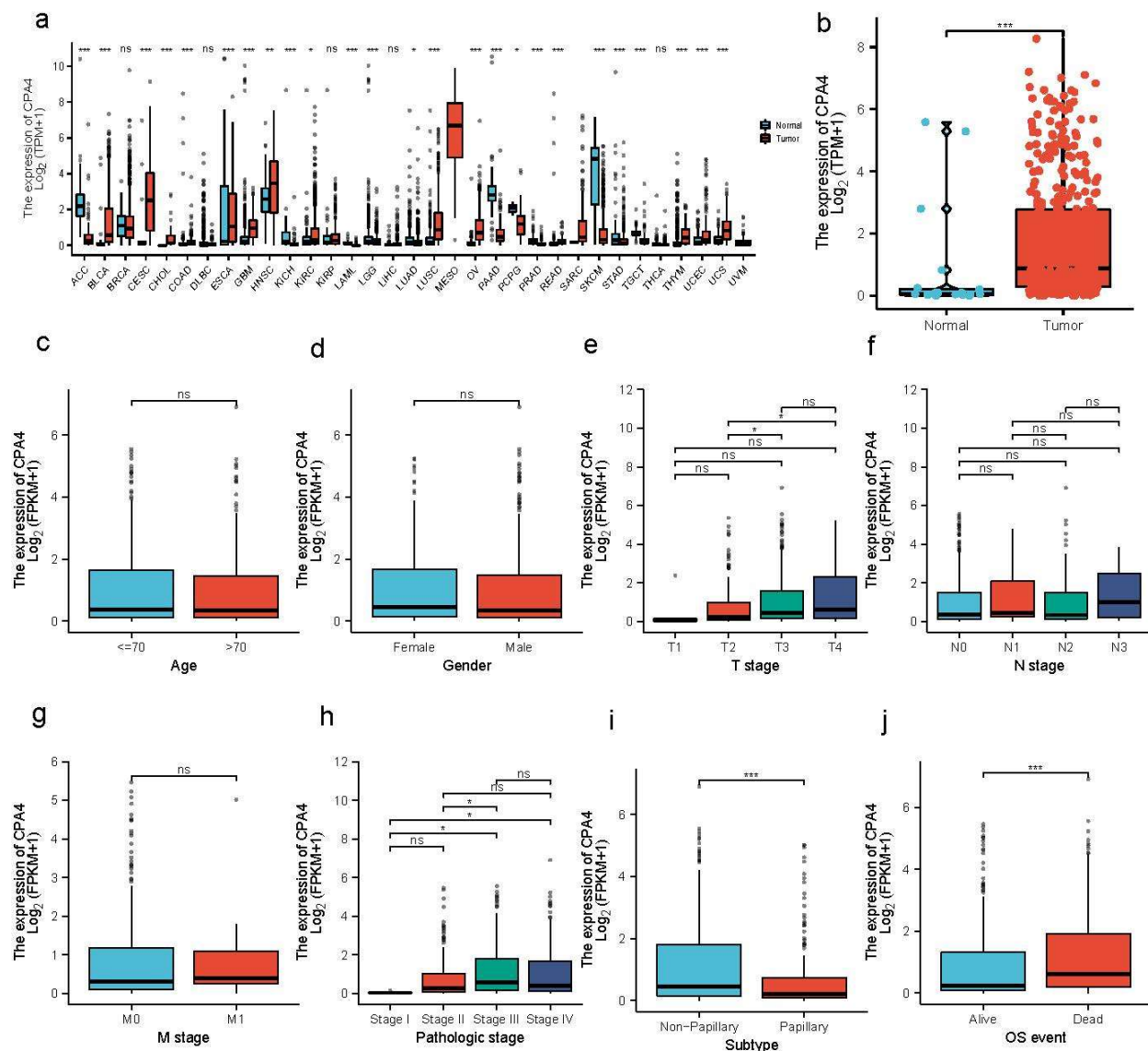
Figure 1. CPA4 expression and clinicopathological features in BLCA.   (**a**) Human CPA4 expression levels in different cancer tissues and corresponding normal tissues. (**b**) The expression level of CPA4 in BLCA tissue was significantly higher comparing with the normal tissues which from TCGA+GTEx. (**c-g**) There were no statistically significant difference between the expression levels of CPA4 in BLCA and age, gender, pathologic T, N, M stage and pathologic. (**h-j**) Higher Dead event, papillary were associated with higher expression of CPA4 in BLCA

3.3 Impact of high CPA4 expression on the detection and prognosis of BLCA patients

The expression of CPA4 indicated significant discriminative power of to identify tumors from the normal with an AUC value of 0.798 (**Figure2 d**). Kaplan-Meier survival analysis showed that BLCA patients with higher CPA4 expression have shorter overall survival, disease specific survival and progress free interval （**Figure2 a-c**）. KM plots show that higher expression of CPA4 had a worse prognosis than lower. Promoter methylation of CPA4 in TCGA-BLCA data was significantly lower than that of normal tissues adjacent to cancer in UALCAN webpage (p < 0.001; **Figure2 e**)
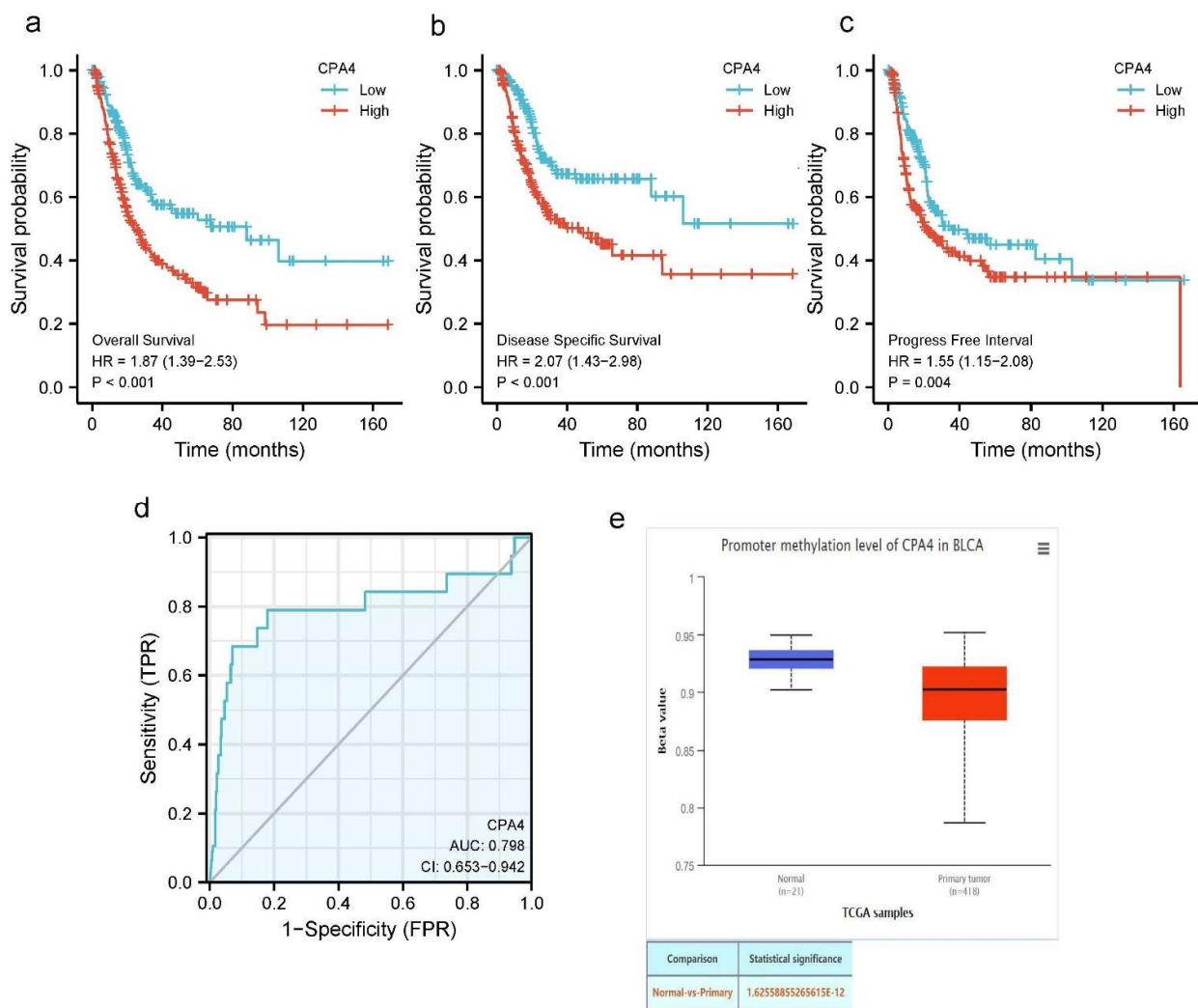
Figure2 (**a-c**) Kaplan-Meier survival curves comparing the high and low expression of CPA4 in BLCA patients. (**a**) Overall survival. (**b**) Disease specific survival. (**c**) Progression-free interval. (**d**) ROC analysis of CPA4 indicates promising discrimination power between tumor and normal tissues. (**e**) The promoter methylation of CPA4 in tumor tissues (n = 418) and normal tissues (n = 21) from TCGA-BLCA data.

3.4 Differentially expressed genes and Go enrichment analysis in high- and low-CPA4 expression samples

We analyzed DEGs in altered expression of CPA4 including the low and high samples to explore the potential mechanisms of CPA4 which promote tumor progression. There were 529 DEGs were identified, of which 349 genes were up regulated, and 180 were down-regulated ($|$log2(FC)$|$ >2 & p.adj <0.05). The DEGs's expression is shown in a heat map and Volcano Plot (**Figure 3**).   By using GO enrichment analysis to predict the co-expression functions in patients with BLCA. The top GO enrichment items in the biological process (BP), molecular function (MF), andcellular component (CC) groups were epidermal cell differentiation, keratinocyte differentiation, keratinization, intermediate filament cytoskeleton, intermediate filament, cornified envelope, endopeptidase inhibitor activity, peptidase inhibitor activity, peptidase inhibitor activity, serine-type endopeptidase inhibitor activity, metabolism of xenobiotics by cytochrome P450, drug metabolism-cytochrome P450 and retinol-metabolism.(**Figure 4a**)
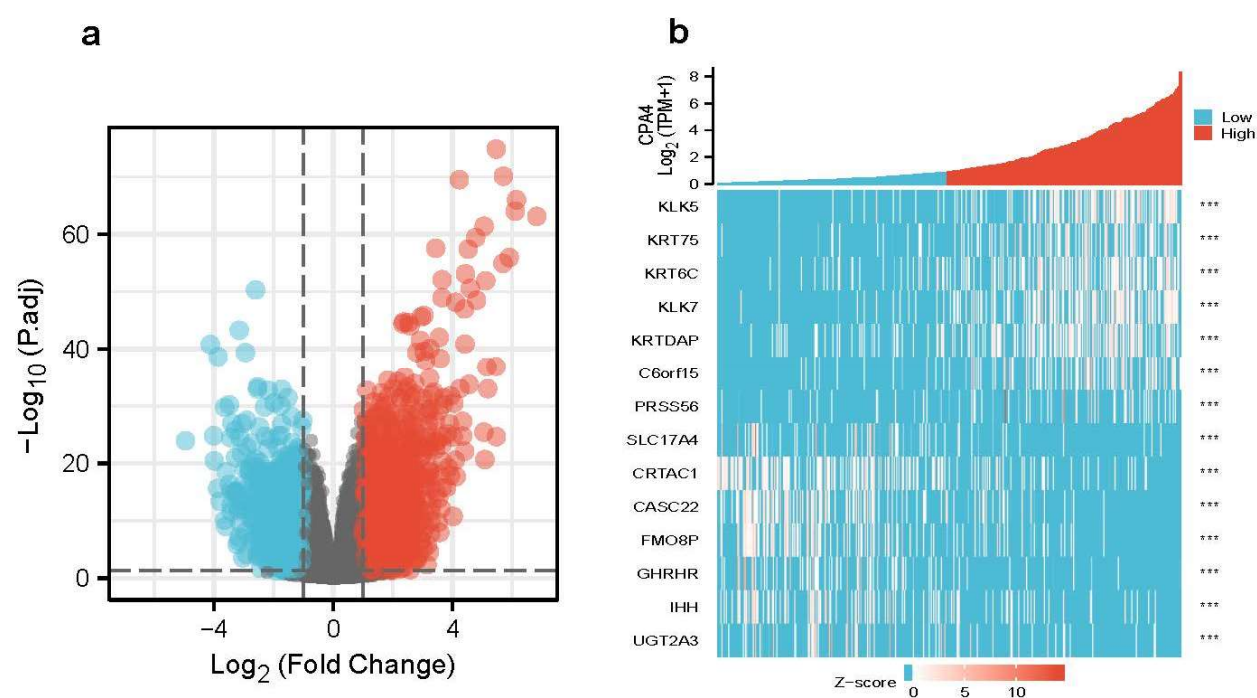
Figure 3 (**a**) Volcano Plot of differentially expressed genes (DEGs). (**b**) Heatmap of differentially expressed genes (DEGs).

3.5 Gene set enrichment analysis for CPA4 related signaling pathways

By the enrichment of MSigDB Collection (c2.all.v7.0.symbols.gmt [Curated]), we used the GSEA to identify signaling pathways associated with CPA4 between the different expression level of CPA4 with significant difference (adjusted P-value < 0.05, FDR < 0.25). Eight pathways including formation of the cornified envelope, keratinization, immunoregulatory interactions between a lymphoid and a non-lymphoid cell, wp hair follicle development cytodifferentiation part 3 of 3, antigen processing and presentation, assembly of collagen fibrils and other multimeric structures, graft versus host disease, cytokine-cytokine receptor interaction (**Figure 4**)
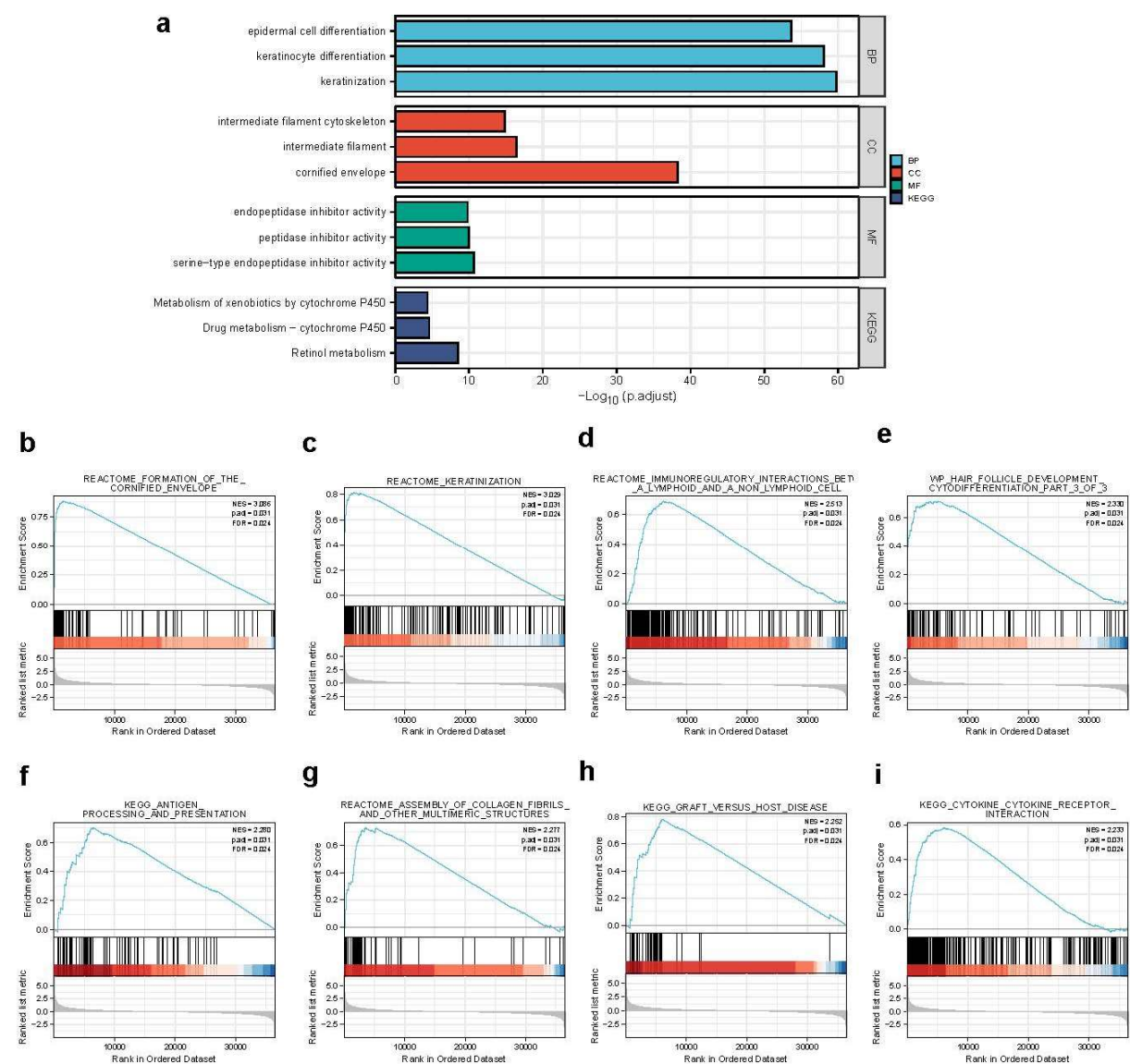
Figure 4 (**a**) GO enrichment analysis of differentially expressed genes (DEGs) in high- and low-CPA4 expression samples. (**b-c**) Enrichment plots from GSEA. Several pathways were differentially enriched in BLCA patients according to different CPA4 expression. (**b**) Formation of the cornified envelope. (**c**) Keratinization. (**d**) Immunoregulatory interactions between a lymphoid and a non-lymphoid cell. **e** WP hair follicle development cytodifferentiation part 3 of 3. (**f**) Antigen processing and presentation. (**g**) Assembly of collagen fibrils and other multimeric structures. (**h**) Graft versus host disease. (**i**) Cytokine-cytokine receptor interaction. ES, enrichment score; NES, normalized enrichment score; ADJ p-Val, adjusted P-value; FDR, false discovery rate.

3.6 CPA4 expression predict poor prognosis in different cancer stages

Univariate cox proportional-hazards model analysis showed high CPA4 expression, high pathologic grade and stage (T, N, M) and subtype of papillary were a negative predictor for OS for BLCA patients. Meanwhile, in multivariate regression analysis, CPA4 expression was an independent factor correlated with OS both in the low expression set and high expression set (p = 0.003) (**Figure 5**).
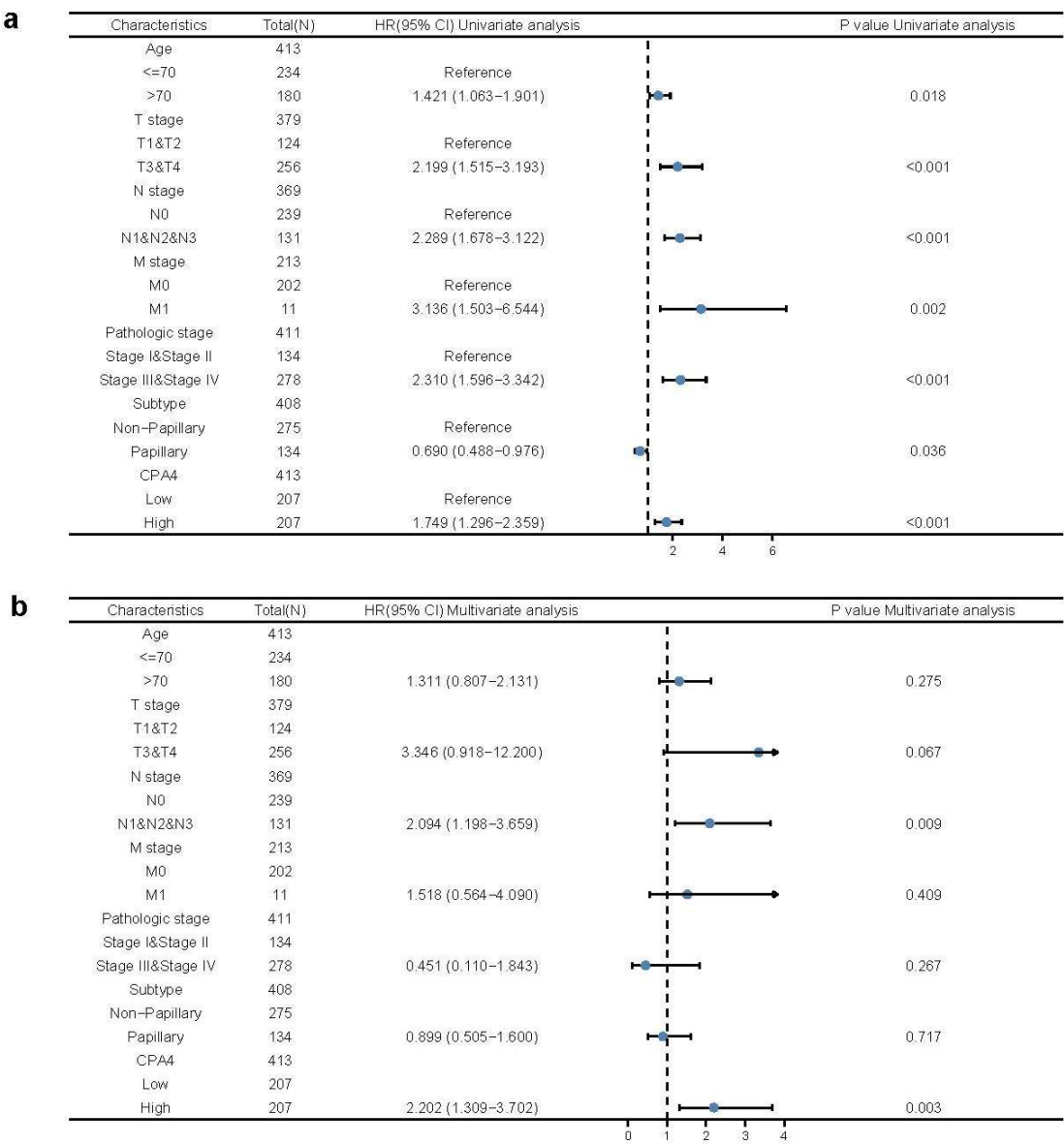
Figure 5 Univariate (**a**) and multivariate (**b**) regression analysis of CPA4 and other clinicopathologic parameters with OS in BLCA patients

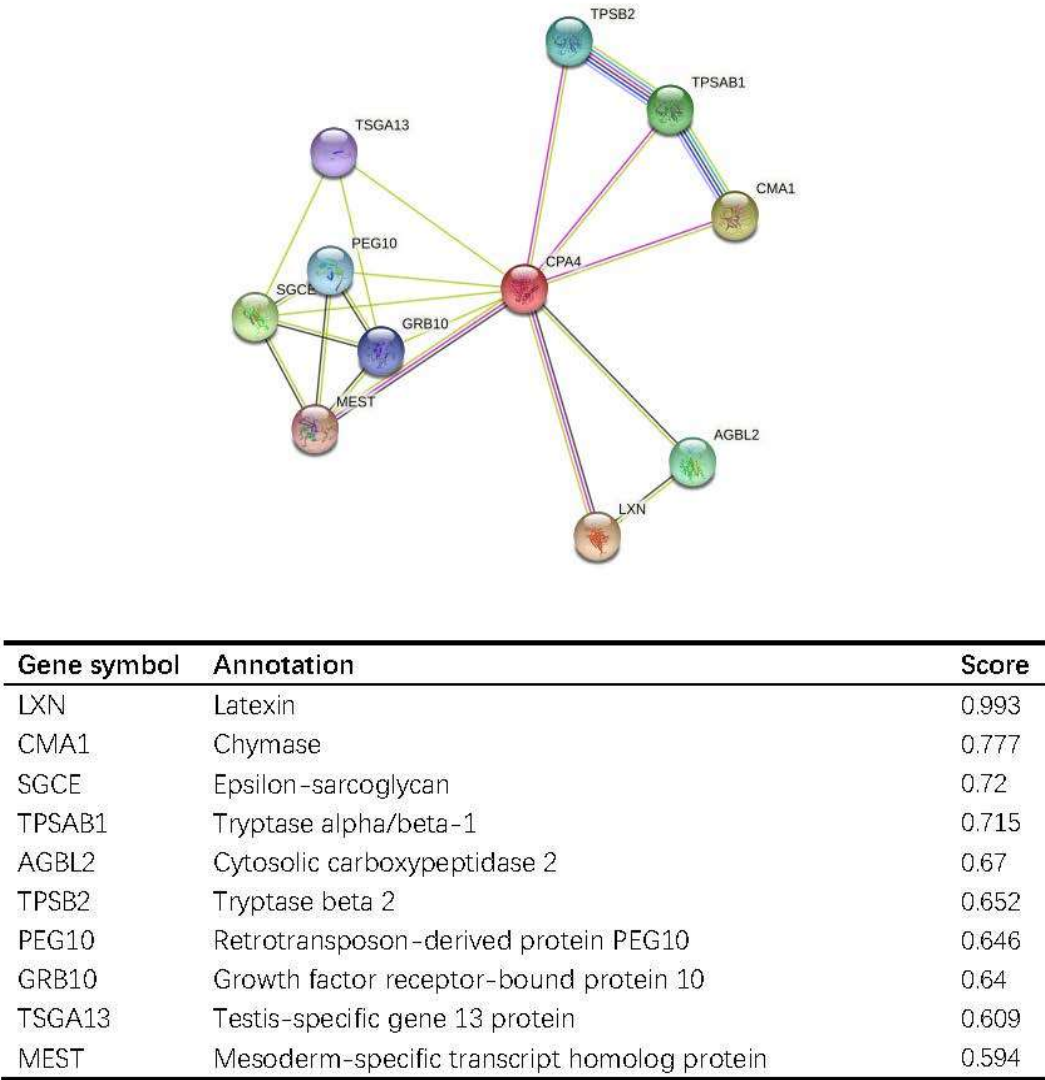*3.7 Construction of nomogram for predicting OS and validation by calibration*

We constructed a nomogram for predicting the prognosis of BLCA with relative clinical situation which integrates the clinical characteristics which associated with the survival of BLCA. Base on multivariate Cox analysis, the nomogram was assigned to the clinical characteristics to a point and the sum of points awarded to each characteristic a point from 0 to 100. All of the points are accumulated and recorded as the total points. Absolute point axis downward to outcome axis determine the probability of BLCA survival at 1, 3, 5-year OS (**Figure 6 a**). Learn from nomogram, the expression of CPA4 contributes great points compared with other relative clinical situations including T, N, M stage, radiation therapy and primary therapy outcome. Meanwhile, calibration plot indicated great agreement between the predicted and observed values which are close to the 45-dgree line as the ideal curve (**Figure 6 b**).
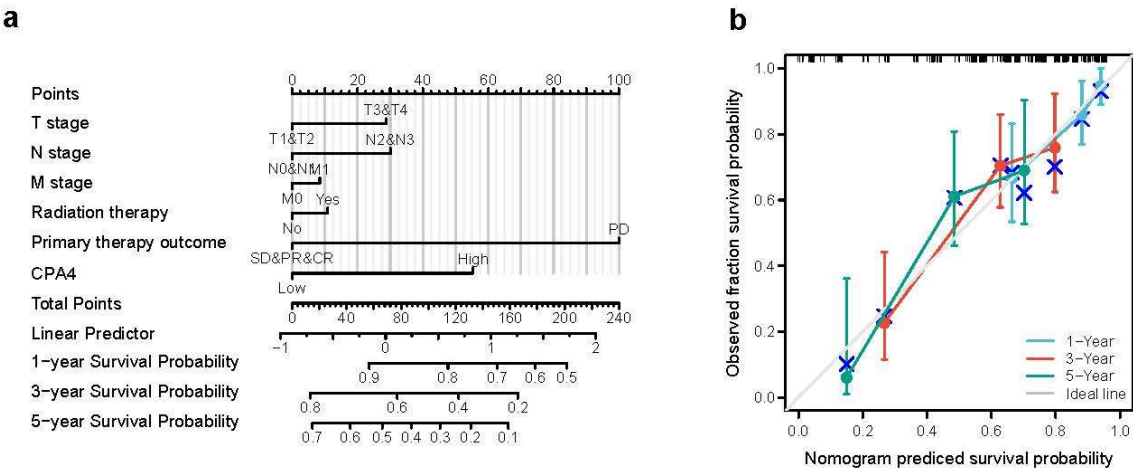
Figure 6 The relationship of CPA4 expression relate to other clinical factors with overall survival (OS). (**a**) Nomogram for predicting the probability of 1-, 3-, and 5-year OS for BLCA patients. (**b**) Calibration plot of the nomogram for predicting the OS likelihood

3.8 CPA4-interaction proteins networks in BLCA tissue

CPA4-interaction proteins networks had been constructed to further explore the necessary proteins for the metabolism and molecular mechanism used by STRING. PPI network of CPA4 protein showed the relationship of the CPA4 protein in the progression of BLCA. There are 10 proteins and corresponding gene names had been listed with their annotations score (Figure 7). The top 10 gene included LXN, CMA1, SGCE, TPSAB1, AGBL2, TPSB2, PEG10, GRB10, TSGA13, MEST which LXN had the highest score.

Gene symbol	Annotation	Score
LXN	Latexin	0.993
CMA1	Chymase	0.777
SGCE	Epsilon-sarcoglycan	0.72
TPSAB1	Tryptase alpha/beta-1	0.715
AGBL2	Cytosolic carboxypeptidase 2	0.67
TPSB2	Tryptase beta 2	0.652
PEG10	Retrotransposon-derived protein PEG10	0.646
GRB10	Growth factor receptor-bound protein 10	0.64
TSGA13	Testis-specific gene 13 protein	0.609
MEST	Mesoderm-specific transcript homolog protein	0.594

Figure 7 CPA4-interaction proteins in BLCA tissue. Annotation of CPA4-interacting proteins and their co-expression scores

3.9 Correlation analysis between CPA4 expression and infiltrating immune cells

The survival of patients with different cancer including BLCA is associated with the tumor infiltrating immune cells. Learn from the result, the expression level of CPA4 had significant correlation with CD8+T cells (r= 0.287, p=2.29e-8), B cells (r=0.218, p=8.65e-10), neutrophils (r=0.196, p=1.76e-4), dendritic cells (r=0.356, p=2.5e-12). P < 0.05 was considered as the difference is of significance (**Figure 8a**). Furthermore, we analysis 24 immune cells including pDC, NK CD56bright cells, DC, Cytotoxic cells, TFH, B cells, CD8 T cells, Th17 cells, Treg, T cells, Mast cells, iDC, NK cells, Tem, aDC, Neutrophils Th1 cells NK CD56dim cells Macrophages Eosinophils Tgd T helper cells Th2 cells Tcm analyzed the correlation between expression of CPA4 and immune infiltration by ssGSEA with Spearman R. Learn from the result, the expression level of CPA4 was negatively correlated with infiltration levels of NK CD56bright cells (P<0.001) and positively correlated with Cytotoxic cells, T cells, NK cells, idc, Tem, Treg, aDC, Neutrophils, NK CD56dim cells, Macrophages, Th2 cells, Th1 cells. (**Figure 8**)
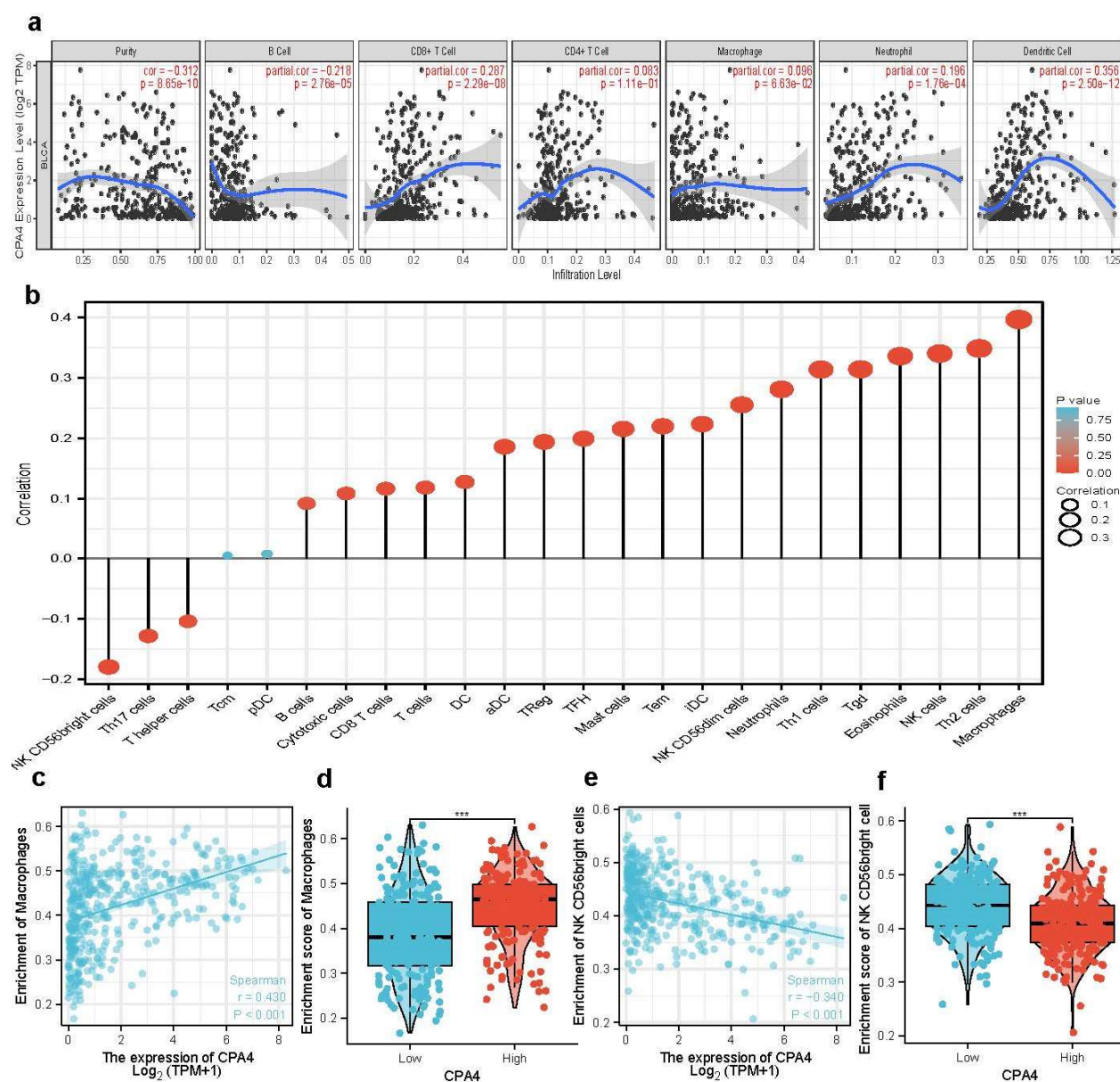
Figure 8 The expression level of CPA4 was related to the immune infiltration in the tumor microenvironment. **b** The forest plot shows the correlation between CPA4 expression level and 24 immune cells. The size of dots indicates the absolute value of Spearman r. **c-d** The Wilcoxon rank sum test was used to analyze the difference of Macrophages cells infiltration level between CPA4 high and low expression groups. **e-f** The correlation between CPA4 expression and NK CD56bright cells infiltration level.

3.10 Possible role of the expression of CPA4 in various infiltrating immune cells

We used the TIMER2 and GEPIA2 to further identify the possible role of the expression of CPA4 in various infiltrating immune cells including, T cells (general), M1/M2 macrophages, tumor-associated macrophages, B cells, neutrophils, monocytes, NK, CD8+ T cells, DCs functional as well as T cells such as Th1, Th2, Th9, Th17, Th22, Tfh, exhausted T cells, and Treg. Learn from the result, Th1, T cell exhaustion, TAM sets marking were great connected with the expression of CPA4 in BLCA (**Table 2**).

Table 2. Correlation analysis between CPA4 and markers of immune cells of BLCA patients in TIMER2 and GEPIA2.

Cell type	Gene marker	None		Purity		Tumor		Normal	
		Cor	P	Cor	P	R	P	R	P
B cell	CD19	-0.042	0.397	-0.138	**	-0.032	0.52	-0.033	0.89
	CD20(KRT20)	-0.314	***	-0.226	***	-0.13	*	-0.18	0.47
	CD38	0.301	***	0.148	**	0.1	*	-0.032	0.9
CD8+ T cell	CD8A	0.267	***	0.12	*	-0.032	0.9	-0.074	0.76
	CD8B	0.15	**	0.018	0.727	0.0031	0.95	-0.096	0.69
Tfh	BCL6	-0.247	***	-0.214	***	-0.14	**	-0.28	0.25
	ICOS	-0.307	***	0.154	**	0.13	**	-0.095	0.7
	CXCR5	0.109	*	-0.095	0.0677	0.075	0.13	0.039	0.87
Th1	T-bet(TBX21)	0.227	***	0.046	0.375	0.2	***	0.041	0.87
	STAT4	0.37	***	0.223	***	0.16	**	0.0068	0.98
	IL12RB2	0.403	***	0.327	***	0.24	***	-0.22	0.37
	WSX1(IL27RA)	0.39	***	0.291	***	0.18	***	0.027	0.91
	STAT1	0.386	***	0.282	***	0.24	***	-0.14	0.56
	IFN-γ(IFNG)	0.278	***	0.161	**	0.13	**	-0.085	0.73
	TNF-α(TNF)	0.287	***	0.194	***	0.098	*	0.27	0.26
Th2	GATA3	-0.484	***	-0.402	***	-0.26	***	-0.26	0.28
	CCR3	0.188	***	0.131	*	0.67	*	-0.14	0.58
	STAT6	-0.228	***	-0.209	***	-0.13	**	-0.32	0.18
	STAT5A	0.004	0.936	-0.158	**	-0.015	0.76	-0.53	*
Th9	TGFBR2	0.087	0.079	-0.014	0.792	0.038	0.45	-0.45	0.056
	IRF4	0.188	***	-0.03	0.571	0.043	0.39	-0.12	0.63
	PU.1(SPI1)	0.356	***	0.181	***	0.15	**	-0.17	0.49
Th17	STAT3	0.325	***	0.232	***	0.15	**	-0.11	0.64
	IL-21R	0.318	***	0.132	*	0.073	0.14	-0.1	0.68
	IL-23R	-0.003	0.945	-0.076	0.143	-0.0048	0.92	-0.019	0.94
	IL-17A	-0.019	0.705	-0.057	0.274	-0.051	0.31	-0.18	0.47
Th22	CCR10	-0.025	0.626	-0.068	0.195	0.029	0.57	-0.34	0.16
	AHR	-0.271	***	-0.195	***	-0.11	*	-0.29	0.23
Treg	FOXP3	0.287	***	0.15	**	0.16	**	0.037	0.88
	CD25(IL2RA)	0.369	***	0.22	***	0.037	0.88	-0.066	0.79
	CCR8	0.218	***	0.083	0.113	0.083	0.094	-0.0059	0.98
T cell exhaustion	PD-1(PDCD1)	0.255	***	0.089	*	0.089	0.073	-0.099	0.69
	CTLA4	0.311	***	0.16	**	0.23	***	-0.11	0.64
	LAG3	0.362	***	0.227	***	0.22	***	-0.19	0.45
	TIM-3(HAVCR2)	0.375	***	0.218	***	0.21	***	-0.097	0.69
Macrophage	CD68	0.316	***	0.193	***	0.14	**	0.49	*
	CD11b(ITGAM)	0.303	***	0.119	*	0.72	*	-0.29	0.23
M1	INOS(NOS2)	-0.033	0.511	-0.092	0.0774	-0.0068	0.89	-0.14	0.57
	IRF5	-0.123	*	-0.116	*	-0.063	0.2	-0.026	0.92
	COX2(PTGS2)	0.209	***	0.164	**	0.057	0.25	-0.24	0.32
M2	CD16	0.408	***	0.273	***	0.17	***	-0.2	0.42
	ARG1	-0.049	0.322	-0.007	0.894	0.076	0.13	0.68	**
	MRC1	0.334	***	0.164	**	0.042	0.4	-0.23	0.34
	MS4A4A	0.353	***	0.199	***	0.12	**	-0.23	0.34
TAM	CCL2	0.26	***	0.113	*	0.022	0.66	-0.12	0.62
	CD80	0.413	***	0.285	***	0.17	***	-0.18	0.46
	CD86	0.396	***	0.244	***	0.17	***	-0.074	0.76
	CCR5	0.29	***	0.101	0.0522	0.13	*	-0.08	0.75
Monocyte	CD14	0.406	***	0.253	***	0.11	*	-0.21	0.38
	CD16(FCGR3B)	0.316	***	0.22	***	0.15	**	-0.073	0.77

	CD115(CSF1R)	0.353	***	0.178	***	0.14	**	-0.28	0.24
Neutrophil	CD66b(CEACAM8)	0.089	0.0721	0.098	0.0609	-0.031	0.53	-0.084	0.73
	CD15(FUT4)	0.141	**	0.047	0.369	0.0041	0.93	-0.33	0.17
	CD11b(ITGAM)	0.303	***	0.119	*	0.018	0.72	-0.29	0.23
Natural killer cell	XCL1	-0.01	0.844	-0.005	0.93	-0.06	0.23	0.13	0.59
	CD7	0.304	***	0.131	*	0.15	**	-0.029	0.91
	KIR3DL1	0.136	**	0.049	0.346	0.075	0.13	0.19	0.44
Dendritic cell	CD1C(BDCA-1)	0.086	0.0823	-0.054	0.305	-0.023	0.65	-0.02	0.93
	CD141(THBD)	0.356	***	0.322	***	0.055	0.27	0.37	0.12
	CD11c(ITGAX)	0.35	***	0.181	***	0.099	*	-0.2	0.41

BLCA Bladder Urothelial Carcinoma; Tfh Follicular helper T cell, Th T helper cell, Treg Regulatory T cell; TAM Tumor-associated macrophage;

None, Correlation without adjustment Correlation; Purity, Correlation adjusted by purity; Tumor, correlation analysis in tumor tissue of TCGA; Normal, correlation analysis in normal tissue of TCGA；

Cor, R value of Spearman's correlation P*< 0.05; P**< 0.01; P*** < 0.001；

4. Discussion

CPA4 (carboxypeptidase A4) is a fellow of the metallocarboxypeptidase family which is a zinc-containing exopeptidase that catalyzes the release of carboxy-terminal amino acids[29]. Learn from other studies, CPA4 is greatly connected to various cancer cells about differentiation, growth and aggressiveness. Furthermore, some studied have indicated that CPA4 as an important regulator which played a key role in the process of inflammation, tumor microenvironment (TME) formation and metastasis which may show an integral part of the adipogenic program like pro-adipogenic effects of FGF-1.[11,12,21] Recently, CPA4 has shown the potential possibility as a biomarker in the early diagnosis for clinical benefit from certain cancers. Currently, the function of CPA4 in tumors had not been fully reported. Despite the potential significance of CPA4 expression in plenty types of cancer, there are no previous studies have ever shown the expression level of CPA4 and clinical significance in bladder cancer.

Based on pan-cancer analysis, our study was demonstrated that human CPA4 expression levels were highly expressed in 11 types of cancer with corresponding normal tissues including BLCA. Our finding agrees with the previous report by Sun and Handa, T. *et al*[17,30,31].

Moreover, previous study has showed that CPA4 expression was detected specifically in the cytoplasm of cancer tissue cells, and in the CPA4-suppressed triple-negative breast cancer (TNBC), viability and migration were decreased.[30] It can act as a potential biomarker of poor prognosis in TNBC. In our study, the results are consistent. We compared the different expression level of BLCA in age, gender, pathologic T, N, M stage, pathologic, subtype and OS. Surprisingly, we found higher dead event, subtype of non-papillary were associated with higher expression of CPA4 in BLCA with statistically significant difference which indicated that CPA4 was also a potential prognostic biomarker, with expression upregulated in tumor tissues. However, Sun et al. showed that the expression of CPA4 in pancreatic cancer was higher than the normal one which contradicts our current results[13], possible reason may contribute to more of the samples analyzed in our study come from the primary tumor rather than metastases. Learn from Kaplan-Meier curves, we confirmed that overexpression of CPA4 is associated with shorter overall survival (OS), disease specific survival (DSF), progress free interval (PFI). Our study explored the mechanism of CPA4 overexpression in BLCA, which showed that CPA4 overexpression might be related to CPA4 methylation. It was Interesting that CPA4 methylation was adverse associated with the prognosis of BLCA. From ROC curve, CPA4 may be a potential diagnostic biomarker between bladder cancer from normal tissues with AUC 0.798. By univariate and multivariate regression analysis, we found that high CPA4

expression, high pathologic grade and stage (T, N, M) and subtype of papillary were a negative predictor for OS in BLCA patients and CPA4 can be independent factor correlated with OS. The nomogram predicted the 1-, 3-, 5-year OS more accurately for patients with BLCA which could help to screen and determine more aggressive treatment for those high-risk patients with BLCA. Pathways of formation of the cornified envelope, keratinization, immunoregulatory interactions between a lymphoid and a non-lymphoid cell, wp hair follicle development cytodifferentiation part 3 of 3, antigen processing and presentation, assembly of collagen fibrils and other multimeric structures, graft versus host disease, cytokine-cytokine receptor interaction are differentially enriched in CPA4 high expression phenotype.

PPI network of CPA4 protein, which were constructed by STRING. showed the relationship of the CPA4 in the progression of BLCA such as LXN, CMA1, SGCE, TPSAB1 and so on. It has been reported that latexin (LXN) can inhibits human CPA4, whose expression is induced in prostate cancer cells after treatment with histone deacetylase inhibitors[32]. The level of CMA1, a key gene, is significantly correlated with gastric cancer prognosis and infiltration level[33]. SGCE Promotes Breast Cancer Stem Cells self-renewal, chemoresistance, and metastasis both in vitro and in vivo by Stabilizing EGFR[33]. Thus, it is speculated that high expression of CPA4 may increase the degree of malignancy of tumors through CPA4 interacting proteins, leading to deterioration of patients' conditions.

Moreover, CPA4 plays a specific role in immune infiltration in bladder cancer. Compellingly, we unraveled that several infiltrating immune cells (Th1cell, Th2 cell, T cell exhaustion, TAM) were correlated with the expression of CPA4 in bladder cancer by using TIMER2 and GEPIA2. Type 1 T helper (Th1) cells produce interferon-gamma[34]. Dual inhibition of STAT1 and STAT3 activation downregulates expression of PD-L1 in cancer cells[35]. T-cell exhaustion is a state of T-cell dysfunction that occurs in many chronic infections and cancers[36]. Scholars have observed that CTLA4 was identified as a crucial negative regulator of the immune system which transmits an inhibitory signal[37].

It is of great significance to features of and pathways in BLCA to exhaustion has crucial implications for the success of checkpoint blockade and adoptive T cell transfer therapies[38]

In conclusion, CPA4 may be a novel and great prognostic biomarker based on the bioinformation analysis in BLCA. We have found the expression of CPA4 is great associate with T cell exhaustion like CTLA4 which may guide checkpoint blockade and adoptive T cell transfer therapies. However, further experiments are needed to conduct to in order to confirm the prognostic value of CPA4 in BLCA and more detail mechanisms to explore.

5. Conclusions

Overexpression of CPA4 was related to the poor prognosis, as well as impaired infiltration of immune cells in BLCA as the disease progressed which may guide the application of checkpoint blockade and the adoption of T cell transfer therapy.

References

1.   Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **2018**, *68*, 394-424, doi:10.3322/caac.21492.

2.   Lenis, A.T.; Lec, P.M.; Chamie, K.; Mshs, M.D. Bladder Cancer: A Review. *Jama* **2020**, *324*, 1980-1991, doi:10.1001/jama.2020.17598.

3.   McConkey, D.J.; Choi, W. Molecular Subtypes of Bladder Cancer. *Current oncology reports* **2018**, *20*, 77, doi:10.1007/s11912-018-0727-5.

4.   Bhanvadia, S.K. Bladder Cancer Survivorship. *Current urology reports* **2018**, *19*, 111, doi:10.1007/s11934-018-0860-6.

5.   Babjuk, M.; Burger, M.; Compérat, E.M.; Gontero, P.; Mostafid, A.H.; Palou, J.; van Rhijn, B.W.G.; Rouprêt, M.; Shariat, S.F.; Sylvester, R.; et al. European Association of Urology Guidelines on Non-muscle-invasive Bladder Cancer (TaT1 and Carcinoma In Situ) - 2019 Update. *European urology* **2019**, *76*, 639-657, doi:10.1016/j.eururo.2019.08.016.

6.   Berdik, C. Unlocking bladder cancer. *Nature* **2017**, *551*, S34-s35, doi:10.1038/551S34a.

7.   James, A.C.; Gore, J.L. The costs of non-muscle invasive bladder cancer. *The Urologic clinics of North America* **2013**, *40*, 261-269, doi:10.1016/j.ucl.2013.01.004.

8.   Abdollah, F.; Gandaglia, G.; Thuret, R.; Schmitges, J.; Tian, Z.; Jeldres, C.; Passoni, N.M.; Briganti, A.; Shariat, S.F.; Perrotte, P.; et al. Incidence, survival and mortality rates of stage-specific bladder cancer in United States: a trend analysis. *Cancer epidemiology* **2013**, *37*, 219-225, doi:10.1016/j.canep.2013.02.002.

9.   Grayson, M. Bladder cancer. *Nature* **2017**, *551*, S33, doi:10.1038/551S33a.

10.   Tanco, S.; Zhang, X.; Morano, C.; Avilés, F.X.; Lorenzo, J.; Fricker, L.D. Characterization of the substrate specificity of human carboxypeptidase A4 and implications for a role in extracellular peptide processing. *The Journal of biological chemistry* **2010**, *285*, 18385-18396, doi:10.1074/jbc.M109.060350.

11.   Ross, P.L.; Cheng, I.; Liu, X.; Cicek, M.S.; Carroll, P.R.; Casey, G.; Witte, J.S. Carboxypeptidase 4 gene variants and early-onset intermediate-to-high risk prostate cancer. *BMC cancer* **2009**, *9*, 69, doi:10.1186/1471-2407-9-69.

12.   Kayashima, T.; Yamasaki, K.; Yamada, T.; Sakai, H.; Miwa, N.; Ohta, T.; Yoshiura, K.; Matsumoto, N.; Nakane, Y.; Kanetake, H.; et al. The novel imprinted carboxypeptidase A4 gene (CPA4) in the 7q32 imprinting domain. *Human genetics* **2003**, *112*, 220-226, doi:10.1007/s00439-002-0891-3.

13.   Sun, L.; Burnett, J.; Guo, C.; Xie, Y.; Pan, J.; Yang, Z.; Ran, Y.; Sun, D. CPA4 is a promising diagnostic serum biomarker for pancreatic cancer. *American journal of cancer research* **2016**, *6*, 91-96.

14.   Hong, W.; Xue, M.; Jiang, J.; Zhang, Y.; Gao, X. Circular RNA circ-CPA4/ let-7 miRNA/PD-L1 axis regulates cell growth, stemness, drug resistance and immune evasion in non-small cell lung cancer (NSCLC). *Journal of experimental & clinical cancer research : CR* **2020**, *39*, 149, doi:10.1186/s13046-020-01648-1.

15.   Shao, Q.; Zhang, Z.; Cao, R.; Zang, H.; Pei, W.; Sun, T. CPA4 Promotes EMT in Pancreatic Cancer via Stimulating PI3K-AKT-mTOR Signaling. *OncoTargets and therapy* **2020**, *13*, 8567-8580, doi:10.2147/ott.S257057.

16.   Huang, H.; Reed, C.P.; Zhang, J.S.; Shridhar, V.; Wang, L.; Smith, D.I. Carboxypeptidase A3 (CPA3): a novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells. *Cancer research* **1999**, *59*, 2981-2988.

17.   Sun, L.; Guo, C.; Burnett, J.; Pan, J.; Yang, Z.; Ran, Y.; Sun, D. Association between expression of Carboxypeptidase 4 and stem cell markers and their clinical significance in liver cancer development. *Journal of Cancer* **2017**, *8*, 111-116, doi:10.7150/jca.17060.

18.   Fu, Y.; Su, L.; Cai, M.; Yao, B.; Xiao, S.; He, Q.; Xu, L.; Yang, L.; Zhao, C.; Wan, T.; et al. Downregulation of CPA4 inhibits non small-cell lung cancer growth by suppressing the AKT/c-MYC pathway. *Molecular carcinogenesis* **2019**, *58*, 2026-2039, doi:10.1002/mc.23095.

19.   Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology (Poznan, Poland)* **2015**, *19*, A68-77, doi:10.5114/wo.2014.47136.

20.   Chandrashekar, D.S.; Bashel, B.; Balasubramanya, S.A.H.; Creighton, C.J.; Ponce-Rodriguez, I.; Chakravarthi, B.; Varambally, S. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia (New York, N.Y.)* **2017**, *19*, 649-658, doi:10.1016/j.neo.2017.05.002.

21.   He, J.; Chen, D.L.; Samocha-Bonet, D.; Gillinder, K.R.; Barclay, J.L.; Magor, G.W.; Perkins, A.C.; Greenfield, J.R.; Yang, G.; Whitehead, J.P. Fibroblast growth factor-1 (FGF-1) promotes adipogenesis by downregulation of carboxypeptidase A4 (CPA4) - a negative regulator of adipogenesis implicated in the modulation of local and systemic insulin sensitivity. *Growth factors (Chur, Switzerland)* **2016**, *34*, 210-216, doi:10.1080/08977194.2017.1285764.

22.   Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **2019**, *10*, 1523, doi:10.1038/s41467-019-09234-6.

23.   Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 15545-15550, doi:10.1073/pnas.0506580102.

24.   Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* **2012**, *16*, 284-287, doi:10.1089/omi.2011.0118.

25.   Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* **2019**, *47*, D607-d613, doi:10.1093/nar/gky1131.

26.   Li, T.; Fu, J.; Zeng, Z.; Cohen, D.; Li, J.; Chen, Q.; Li, B.; Liu, X.S. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic acids research* **2020**, *48*, W509-w514, doi:10.1093/nar/gkaa407.

27.   Hänzelmann, S.; Castelo, R.; Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC bioinformatics* **2013**, *14*, 7, doi:10.1186/1471-2105-14-7.

28.   Li, C.; Tang, Z.; Zhang, W.; Ye, Z.; Liu, F. GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA. *Nucleic acids research* **2021**, *49*, W242-w246, doi:10.1093/nar/gkab418.

29.   Wei, S.; Segura, S.; Vendrell, J.; Aviles, F.X.; Lanoue, E.; Day, R.; Feng, Y.; Fricker, L.D. Identification and characterization of three members of the human metallocarboxypeptidase gene family. *The Journal of biological chemistry* **2002**, *277*, 14954-14964, doi:10.1074/jbc.M112254200.

30.   Handa, T.; Katayama, A.; Yokobori, T.; Yamane, A.; Fujii, T.; Obayashi, S.; Kurozumi, S.; Kawabata-Iwakawa, R.; Gombodorj, N.; Nishiyama, M.; et al. Carboxypeptidase A4 accumulation is associated with an aggressive phenotype and poor prognosis in triple-negative breast cancer. *International journal of oncology* **2019**, *54*, 833-844, doi:10.3892/ijo.2019.4675.

31.   Sun, L.; Wang, Y.; Yuan, H.; Burnett, J.; Pan, J.; Yang, Z.; Ran, Y.; Myers, I.; Sun, D. CPA4 is a Novel Diagnostic and Prognostic Marker for Human Non-Small-Cell Lung Cancer. *Journal of Cancer* **2016**, *7*, 1197-1204, doi:10.7150/jca.15209.

32.   Pallarès, I.; Bonet, R.; García-Castellanos, R.; Ventura, S.; Avilés, F.X.; Vendrell, J.; Gomis-Rüth, F.X. Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 3978-3983, doi:10.1073/pnas.0500678102.

33.   Shi, S.; Ye, S.; Mao, J.; Ru, Y.; Lu, Y.; Wu, X.; Xu, M.; Zhu, T.; Wang, Y.; Chen, Y.; et al. CMA1 is potent prognostic marker and associates with immune infiltration in gastric cancer. *Autoimmunity* **2020**, *53*, 210-217, doi:10.1080/08916934.2020.1735371.

34.   Romagnani, S. Th1/Th2 cells. *Inflammatory bowel diseases* **1999**, *5*, 285-294, doi:10.1097/00054725-199911000-00009.

35.   Sasidharan Nair, V.; Toor, S.M.; Ali, B.R.; Elkord, E. Dual inhibition of STAT1 and STAT3 activation downregulates expression of PD-L1 in human breast cancer cells. *Expert opinion on therapeutic targets* **2018**, *22*, 547-557, doi:10.1080/14728222.2018.1471137.

36.   Wherry, E.J. T cell exhaustion. *Nature immunology* **2011**, *12*, 492-499, doi:10.1038/ni.2035.

37.   Walker, L.S.; Sansom, D.M. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. *Nature reviews. Immunology* **2011**, *11*, 852-863, doi:10.1038/nri3108.

38.   Blank, C.U.; Haining, W.N.; Held, W.; Hogan, P.G.; Kallies, A.; Lugli, E.; Lynn, R.C.; Philip, M.; Rao, A.; Restifo, N.P.; et al. Defining 'T cell exhaustion'. *Nature reviews. Immunology* **2019**, *19*, 665-674, doi:10.1038/s41577-019-0221-9.