Hindawi

Research Article

Characterizing the Crosstalk of NCAPG with Tumor Microenvironment and Tumor Stemness in Stomach Adenocarcinoma

Zheng Xiang,[1] **Genlan Cha,**[2] **Yihao Wang,**[1] **Jikai Gao,**[1] **and Jianguang Jia** [1]

[1]*Department of Oncology, The First Affiliated Hospital of Bengbu Medical College, Bengbu, China 233000*
[2]*Department of Radiotherapy, The First Affiliated Hospital of Bengbu Medical College, Bengbu, China 233000*

Correspondence should be addressed to Jianguang Jia; jiajianguang1978@126.com

Background. Nonstructural maintenance of non-SMC condensin I complex subunit G (NCAPG) exerts critical effects on cancer progression. However, its biological roles in tumorigenesis and metastasis remain unclear. Thus, we aimed to assess the prognostic utility of NCAPG in stomach adenocarcinoma (STAD) and its potential as a tumor biomarker. *Methods.* Pan-cancer expression profile dataset from public databases and corresponding clinical information were extracted. Single-sample gene set enrichment analysis (ssGSEA) was performed for the evaluation of immune correlations pan-cancer. Subsequently, we focused on STAD and evaluated the methylation profiles, copy number variants (CNVs), and single nucleotide variants (SNVs). Immune features were analyzed between high and low NCAPG expression groups. Differential analysis was performed between high and low expression groups to identify differentially expressed genes (DEGs). Prognostic DEGs were screened by univariate analysis, and an NCAPG-based risk model was constructed based on the prognostic DEGs and LASSO analysis. *Results.* NCAPG expression in STAD was significantly and positively correlated with four immune checkpoints, namely, CTLA4, PDCD1, LAG3, and CD276, but was negatively correlated with the infiltration of most immune cells. High and low NCAPG expression groups had differential overall survival, tumor mutation burden, and differential enrichment of therapeutic-related pathways. An immune risk scoring model related to NCAPG expression and immune score was constructed which showed a favorable performance in predicting STAD prognosis as well as predicting the response to immunotherapy. In addition, we found a higher mRNA stemness index (mRNAsi) in the high-risk group and a positive correlation between NCAPG expression and mRNAsi. *Conclusion.* NCAPG was suggested to be involved in the regulation of tumor microenvironment in STAD. High NCAPG expression was related to high tumor stemness and good prognosis. The immune risk model had a potential to predict STAD prognosis and help directing therapeutic treatment.

1. Introduction

Cancer is a major reason for the deterioration of quality of life globally and is the top cause of mortality; to date, there is no cure for cancer [1]. Immunotherapy has emerged as a prominent treatment option for cancer in recent years; in particular, immune checkpoint blockade agents are promising [2]. Using improved and updated datasets available through public databases such as The Cancer Genome Atlas (TCGA), new immunotherapeutic targets can be identified by performing pan-cancer gene expression analysis and eval-

uating their correlation with relevant signaling pathways and clinical prognosis [3].

The possible regulatory role of cohesin complexes (condensin complexes) in several cancer types including colon adenocarcinoma and hepatocellular carcinoma has been recently described. In cancer, these condensin complexes exert various impacts on the progression and drug resistance [4]. Nonstructural maintenance of non-SMC condensin I complex subunit G (NCAPG) is essential for chromatin segregation [5]. NCAPG is encoded by the NY-MEL-3 gene and is mapped to chromosome 4p 15.32 in humans [6]. In

prostate cancer, a high NCAPG expression is related to a poor prognosis [7], and its knockdown combined with temozolomide treatment results in coinhibitory impacts on advanced pediatric glioma cells [8]. NCAPG is a hub protein, showing high connectivity in the protein-protein interaction (PPI) network in hepatocellular carcinoma and promotes cellular proliferation [9, 10]. Additionally, studies on stomach adenocarcinoma (STAD) suggest that through the Wnt/β-catenin signaling pathway by promoting epithelial-mesenchymal transition, NCAPG overexpression inhibits apoptosis in cardia adenocarcinoma [11]. Moreover, dysregulation of KNL1, miR-1179, NCAPG, miR-193b-3p, and miR-148a-3p may lead to STAD progression [12]. However, the mechanism underlying NCAPG-mediated proliferation in these cancers remains unknown, and its utility as a tumor-associated biomarker warrants further investigation.

The potential impact of NCAPG on several cancer types warrants clarification. We used two databases, Gene Expression Omnibus (GEO) and TCGA, to assess NCAPG expression in different types of cancers and their relationship with prognosis. We also investigated the immune correlation with NCAPG expression for 33 cancer types and subsequently found the strongest relationship for STAD. Next, we analyzed the potential associations between NCAPG and mutational types, DNA methylation profile, tumor mutation burden (TMB), immune infiltration levels, and clinical responses. We performed a PPI analysis of immune-associated differential genes with NCAPG to investigate its biological functions in the tumor. Finally, an immune risk score (IRS) model was constructed and the performance of the related score was better than that of the TIDE score. In conclusion, our results suggested that NCAPG may underlie implications as a prognostic factor and a potential biomarker for STAD. Thus, we provide insight into the role of NCAPG in STAD immunotherapy.

2. Methods

2.1. Data Source and Preprocessing.
The bioinformatics analysis in this study was partially supported by Sangerbox (http://vip.sangerbox.com/) [13]. Thirty-three TCGA pan-cancer datasets comprising RNA sequencing (RNA-seq) expression profiles, survival information, and somatic mutational profile were extracted from https://xenabrowser.net/ (UCSC Xena). The RNA-seq data in fragments per kilobase million (FPKM) were converted to the transcripts per million (TPM) format, following which $\log_2$ transformations were performed. The downloaded somatic mutation data were already processed by mutect analysis. Finally, following the processing of the copy number variation (CNV) data by the GISTIC algorithm, these were extracted from UCSC Xena, while the methylation profiles were obtained from the LinkedOmics data portal (http://linkedomics.org).

We obtained the STAD-GEO cohort from https://www.ncbi.nlm.nih.gov/geo/ (GEO database), i.e., GSE66229, GSE84437, GSE26942, GSE13861, GSE28541, and GSE26253 comprising detailed survival data. The STAD sample data were preserved.

Three immunotherapy-associated cohorts, namely, GSE135222 (NSCLC), GSE91061 (melanoma), and GSE78220 (melanoma), were extracted. Complete expression data and corresponding detailed clinical information for a STAD immunotherapy-associated cohort (IMvigor210) were obtained using the Creative Commons 3.0 license (from http://research-pub.Gene.com/imvigor210corebiologies/).

2.2. Tumor Immune Microenvironment and Immune Infiltration Analysis.
The single-sample gene set enrichment analysis (ssGSEA) in GSVA R package [14] was used to deal with the problem that single sample cannot perform GSEA. Therefore, the extent of multiple immune cell infiltration was estimated with the ssGSEA function in the GSVA package in R software. Using the Estimation of STromal and Immune cells in MAlignant Tumor tissues using the Expression data method (ESTIMATE), the tumor immune microenvironment score of samples was assessed [15]. Stromal, Immune, and ESTIMATE scores were computed and the differential distributions were compared.

2.3. Functional Enrichment Analysis.
The limma package was utilized for differential expression analysis [16], and differentially expressed genes (DEGs) were identified based on false discovery rate (FDR) < 0.05 with $|\log_2(\text{fold change})| > 1$.

Gene set containing the GSEA input file comprising the expression profile data was obtained (c2.cp.kegg.v7.0.symbols.gmt). Enriched pathways were selected, and GO annotation for DEGs was performed using the WebGestaltR package in R (v0.4.2) with $P < 0.05$ considered statistically significant [17].

2.4. Construction of the PPI Network.
STRING (web version v11.0; https://string-db.org/) database was queried for the known and predicted PPIs. It is a database comprising information from 2031 species and 9.6 million proteins with 13.8 million interactions. The PPI network was constructed to uncover the core regulatory genes and visualized by using Cytoscape.

2.5. Construction of the IRS Model.
In the GSE66229 dataset, random sampling was performed with a sampling ratio of train : test = 1 : 1. Univariate analysis was performed in the training dataset, and prognostic genes were obtained ($P < 0.05$). The best genes were further selected using LASSO. The R package, glmnet [18], was utilized to build the LASSO regression model with prognostic genes as the input.

Based on the formula defined by the sample risk score, the risk-related prognostic risk score (RiskScore) was computed for each sample as follows: RiskScore = $\Sigma \beta_i \times \text{Exp}_i$, where Exp_i indicates the gene signature expression and β is the corresponding gene LASSO regression coefficient. According to the median, patients were classified into high and low RiskScore groups. Survival curves were plotted for prognostic analysis by the Kaplan-Meier method, and the significance of differences was determined using a log-rank test. ROC analysis of RiskScore was performed using the timeROC package [19] in R for prognostic classification.
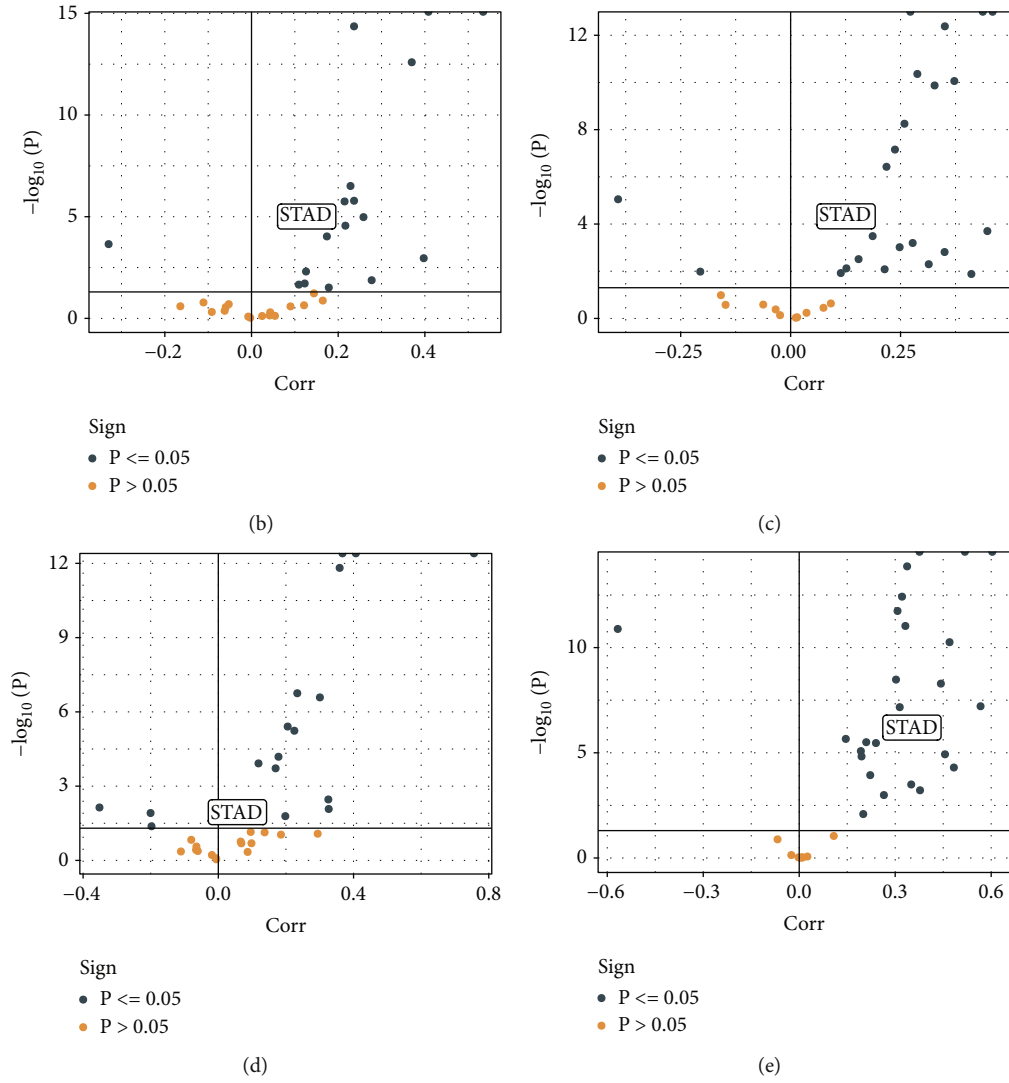
Figure 1: Continued.
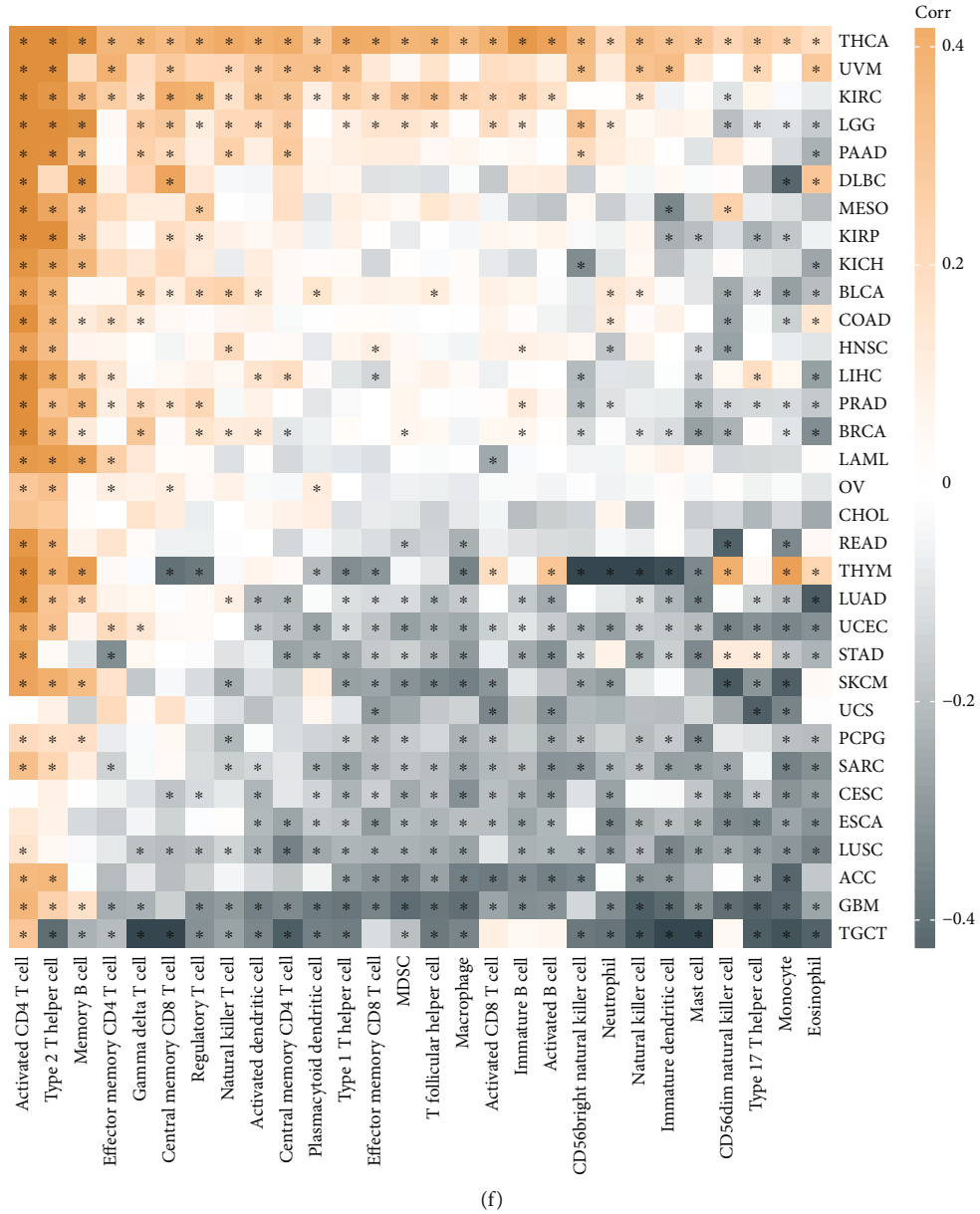
(b)

(c)

(d)

(e)
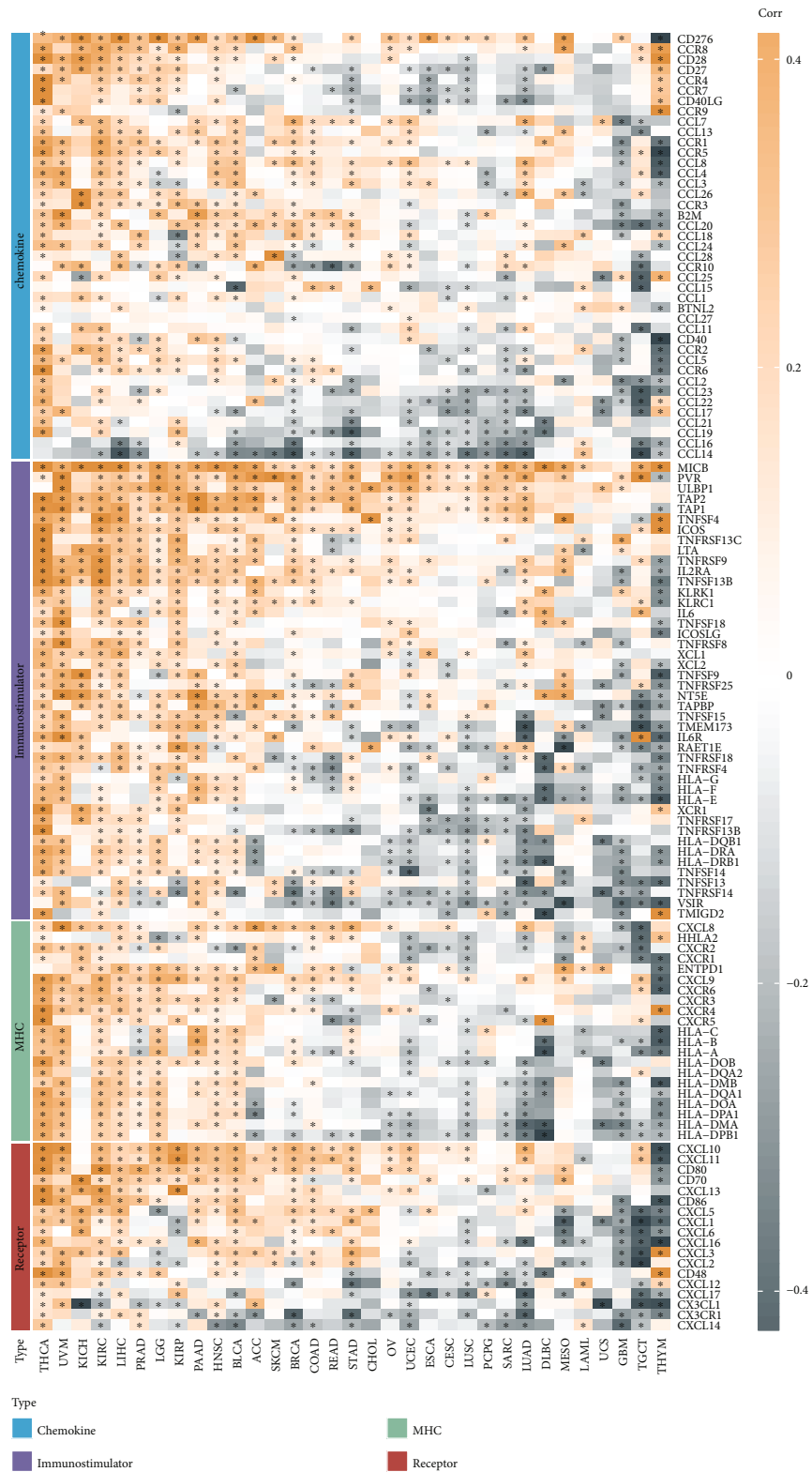
Figure 1: Continued.

(f)

FIGURE 1: Pan-cancer immunological correlation with gene NCAPG. (a) Correlation between NCAPG and immunomodulators (chemokines, receptors, MHC, and immunostimulants). (b–e) Correlation between NCAPG and four immune checkpoints, namely, PDCD1, CTLA4, CD274, and LAG3. Points represent the cancer types. The y-axis represents the Pearson correlation coefficient, while the x-axis represents $-\log_{10}$ (P value). (f) Spearman correlation coefficients between NCAPG gene expression and 28 tumor-associated immune cells in 33 different cancer types are represented by colors. Asterisks indicate statistically significant P values for Spearman correlation analysis (* represents $P < 0.05$).

2.6. TIDE Analysis for the Effectiveness of Immunotherapy. Tumor Immune Dysfunction and Exclusion (TIDE) [20] analysis is utilized to identify biomarkers to predict the therapeutic responses to immune checkpoint inhibitors/drugs by a comprehensive analysis of several tumor expression profiles (http://tide.dfci.harvard.edu/). The sensitivity to immune checkpoint inhibitors is obtained as a TIDE score.

2.7. Calculation of Tumor Stemness. The expression data of embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) was downloaded from the Progenitor Cell Biology Consortium (PCBC) database (https://www.synapse.org/pcbc) [21]. The Ensembl IDs of SC samples were transferred to gene symbols. Only the protein genes were remained. A total of 78 SC samples were included. Gelnet (v1.2.1) R package (https://cran.r-project.org/web/packages/gelnet/index.html) was applied to analyze the mRNA stemness index (mRNAsi) of SCs. For the expression data, mean centering was used to normalize the samples. One-class logistic regression (OCLR) was conducted to calculate the weight vector of each gene. Then, we predicted the mRNAsi
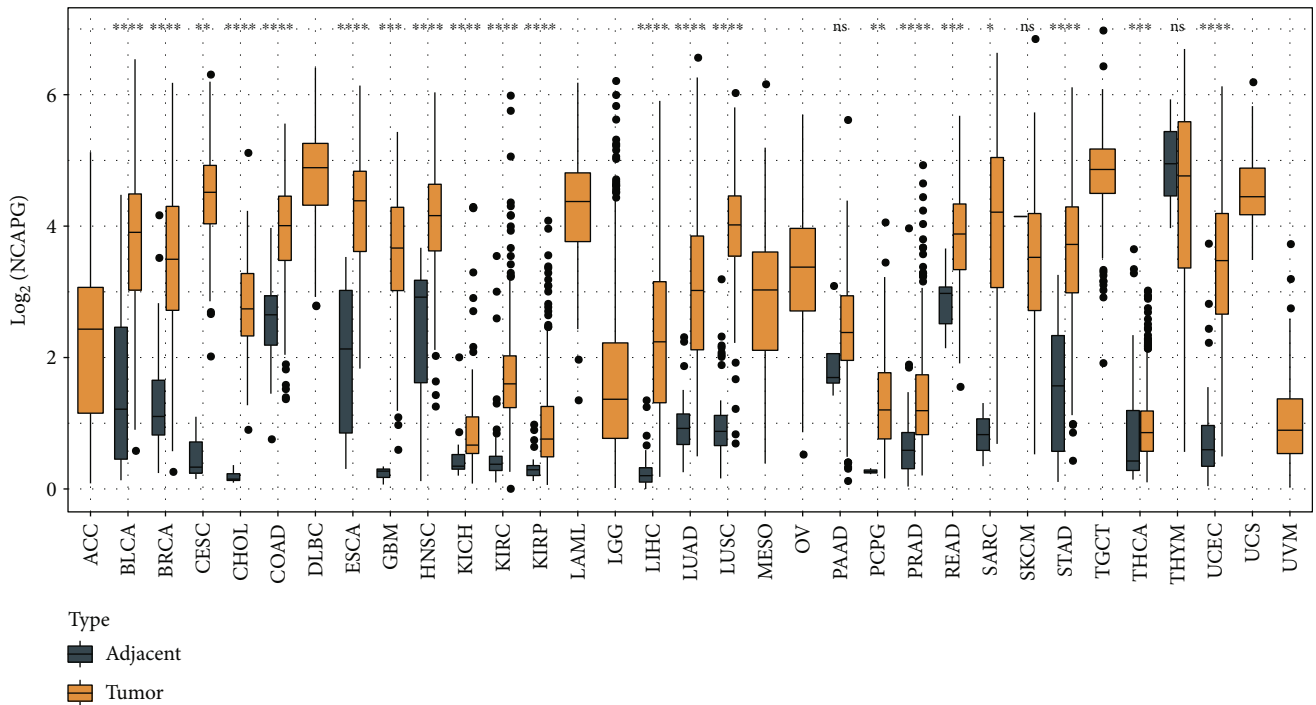
FIGURE 2: Pan-cancer differential expression of gene NCAPG between cancer and para-cancer tissues. Yellow represents significantly high expression in tumors (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, $P > 0.05$).

of TCGA samples based on the Spearman correlation analysis between the expression data and the weight vector of genes. The Spearman correlation coefficients (coef) were normalized according to the formula $mRNAsi = [coef - min (coef)/max (coef)]$, where min (coef) represents the minimum coefficient and max (coef) represents the maximum coefficient. Finally, mRNAsi ranging from 0 to 1 was obtained.

3. Results

3.1. Pan-Cancer Immunological Correlation with NCAPG. Four types of immunological characteristics, namely, chemokines, immunostimulators, MHCs, and receptors, were identified from the literature [22]. The Spearman correlation coefficient was computed between NCAPG and the genes of immunological characteristics in pan-cancer data, and the correlations varied significantly across different cancer types, such as UVM (uveal melanoma), KICH (kidney chromophobe), and THCA (thyroid carcinoma) tumors showing positive correlation, while THYM (thymoma), TGCT (testicular germ cell tumors), and others exhibiting negative correlation (Figure 1(a)). We evaluated the correlation of four key immune checkpoint genes, namely, PDCD1, CTLA4, CD276, and LAG3, with NCAPG in different tumors by Pearson correlation analysis. The result exhibited that they were all positively correlated with NCAPG expression in STAD (Figures 1(b)–1(e)). We evaluated 28 immune cell scores in different cancer types by ssGSEA and calculated their correlation with NCAPG. NCAPG gene expression in STAD showed a significant negative correlation with the

immune scores of 16 immune cells (Figure 1(f)). We assessed the gene expression of NCAPG in pan-cancer and found significant differences of NCAPG expression between tumor samples and adjacent samples (Figure 2). NCAPG expression was elevated in 21 tumor samples, including HNSC (head and neck squamous cell carcinoma), CHOL (cholangiocarcinoma, bile duct cancer), CESC (cervical squamous cell carcinoma and endocervical adenocarcinoma), BLCA (bladder carcinoma), ESCA (esophageal carcinoma, esophageal cancer), COAD (colon adenocarcinoma, colon cancer), BRCA (breast invasive carcinoma), KICH (kidney renal clear cell carcinoma), GBM (glioblastoma multiforme), LIHC (liver hepatocellular carcinoma), KIRC (kidney renal clear cell carcinoma), KIRP (kidney renal papillary cell carcinoma), LUSC (lung squamous cell carcinoma), LUAD (lung adenocarcinoma), PRAD (prostate adenocarcinoma), PCPG (pheochromocytoma and paraganglioma), SARC (sarcoma), UCEC (uterine corpus endometrial carcinoma), READ (rectum adenocarcinoma, rectal adenocarcinoma), STAD, and THCA.

According to the results of pan-cancer analysis, NCAPG correlated with the four types of immunological signatures as well as four important immune checkpoints in STAD. The expression of NCAPG was significantly lower in para-cancer relative to STAD samples. Furthermore, survival time between low and high expression groups showed significant differences in most of the tumors (Fig. S1). Therefore, we were interested in elucidating the role of NCAPG in STAD.

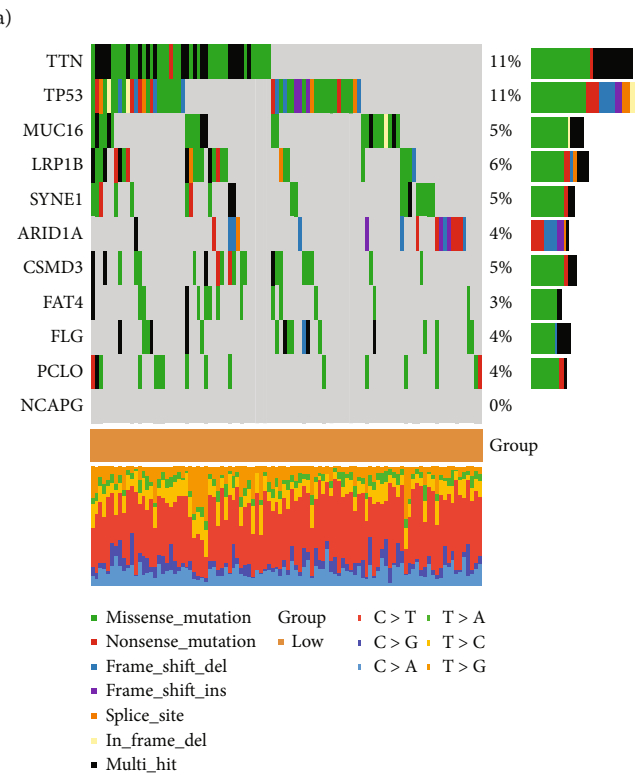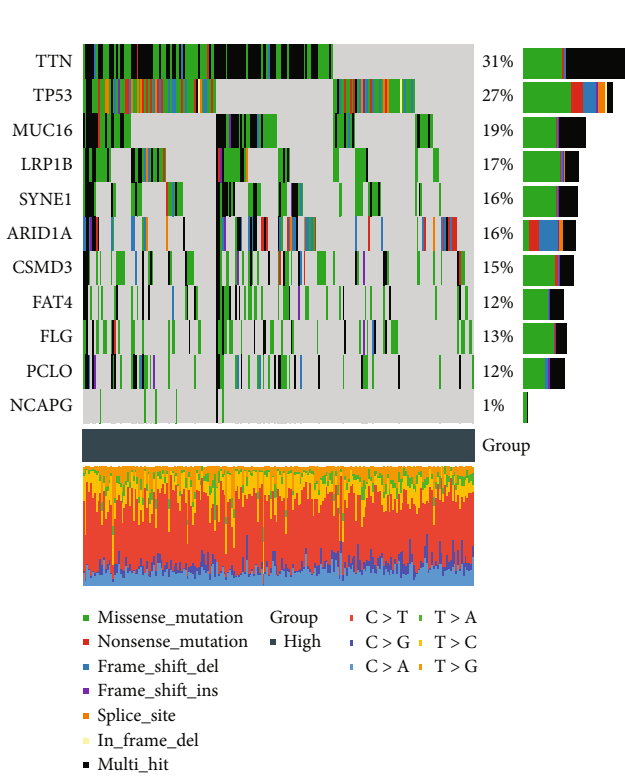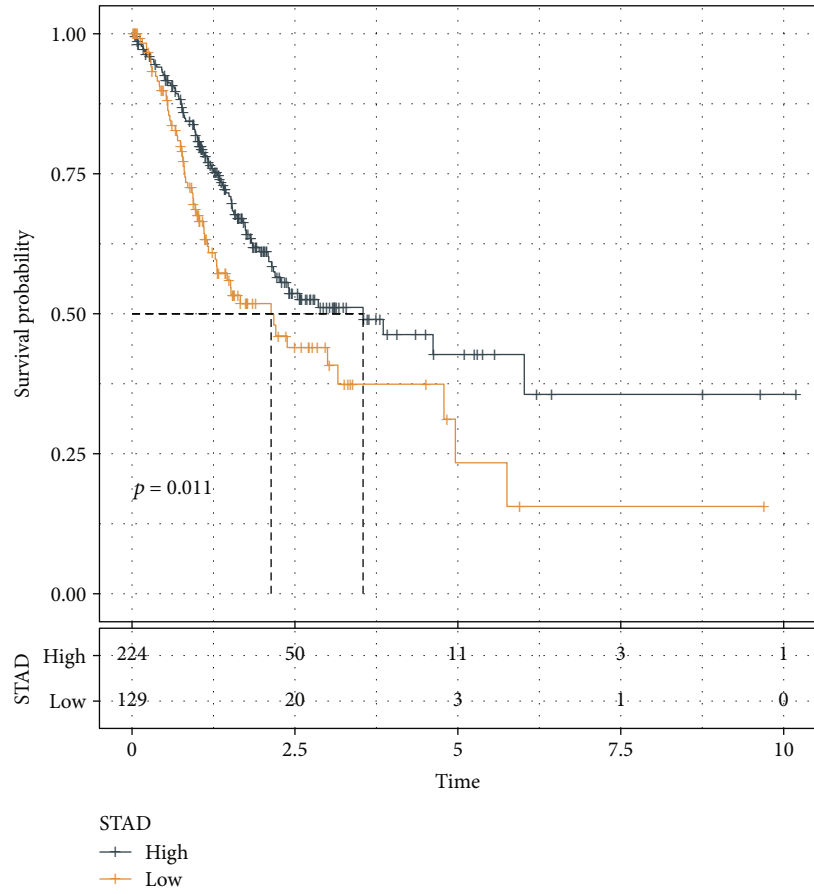*3.2. Analysis of SNVs, CNVs, and Methylation Profile in STAD.* The STAD samples were classified into two groups

(a)
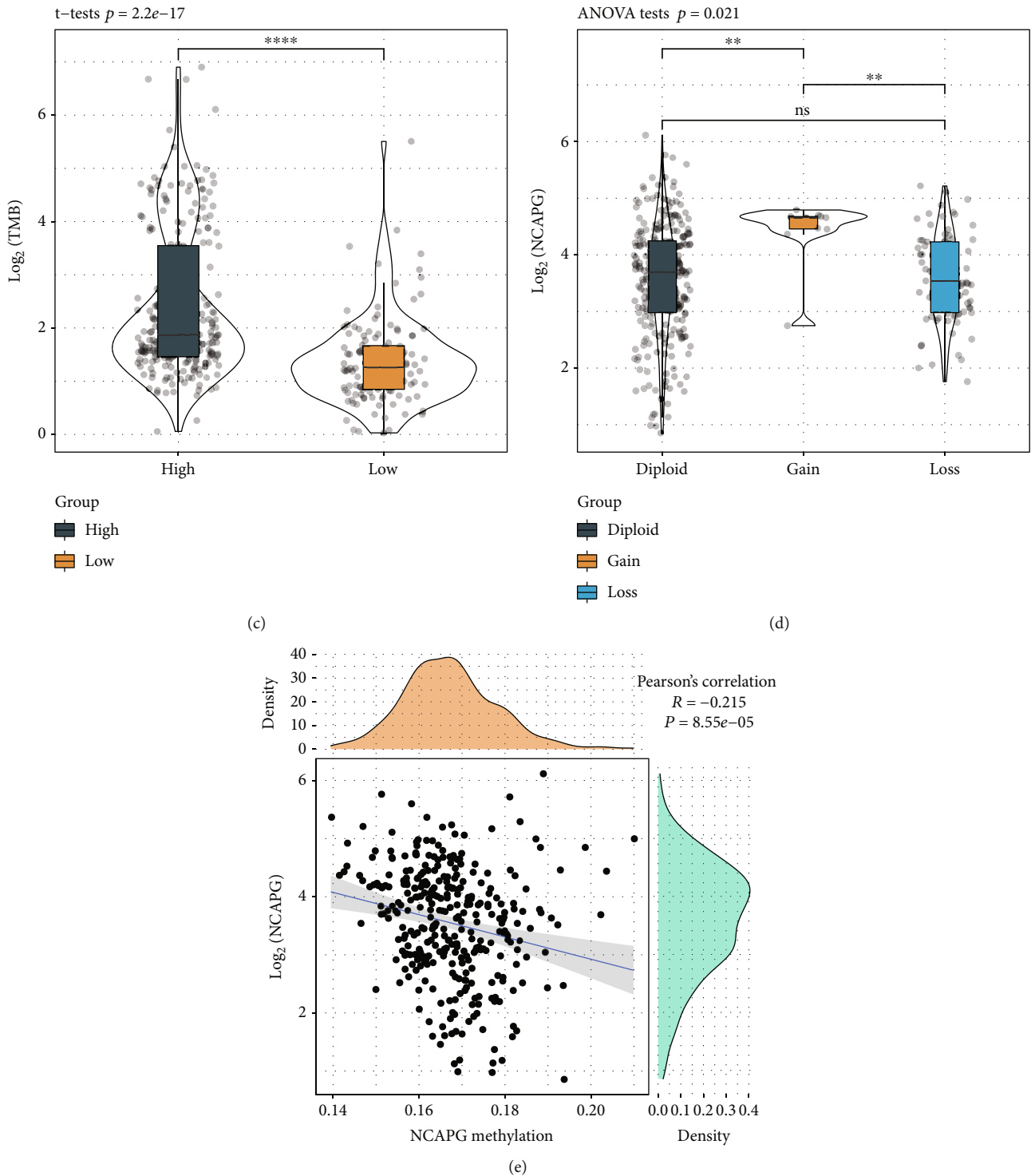

(b)

Figure 3: Continued.

(c)

(d)

(e)

FIGURE 3: Analysis of SNVs, CNVs, and methylation profiles in STAD. (a) KM curve with 0 as the cutoff after expression z-score transformation; (b) mutational distribution for the top 10 genes with the highest mutation frequency in NCAPG expression subgroups; (c) comparison of TMB distribution between NCAPG expression subgroups; (d) NCAPG amplification subgroups based on expression differences; (e) correlation analysis for NCAPG expression and methylation status (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, $P > 0.05$).

with the best cutoff for NCAPG expression determined by survminer R package (Figure 3(a)). The high NCAPG expression group showed a more favorable prognosis, indi-

cating that NCAPG may be a protective factor for STAD. Next, we plotted the 10 genes with the highest mutation frequencies between the expression groups. The mutation
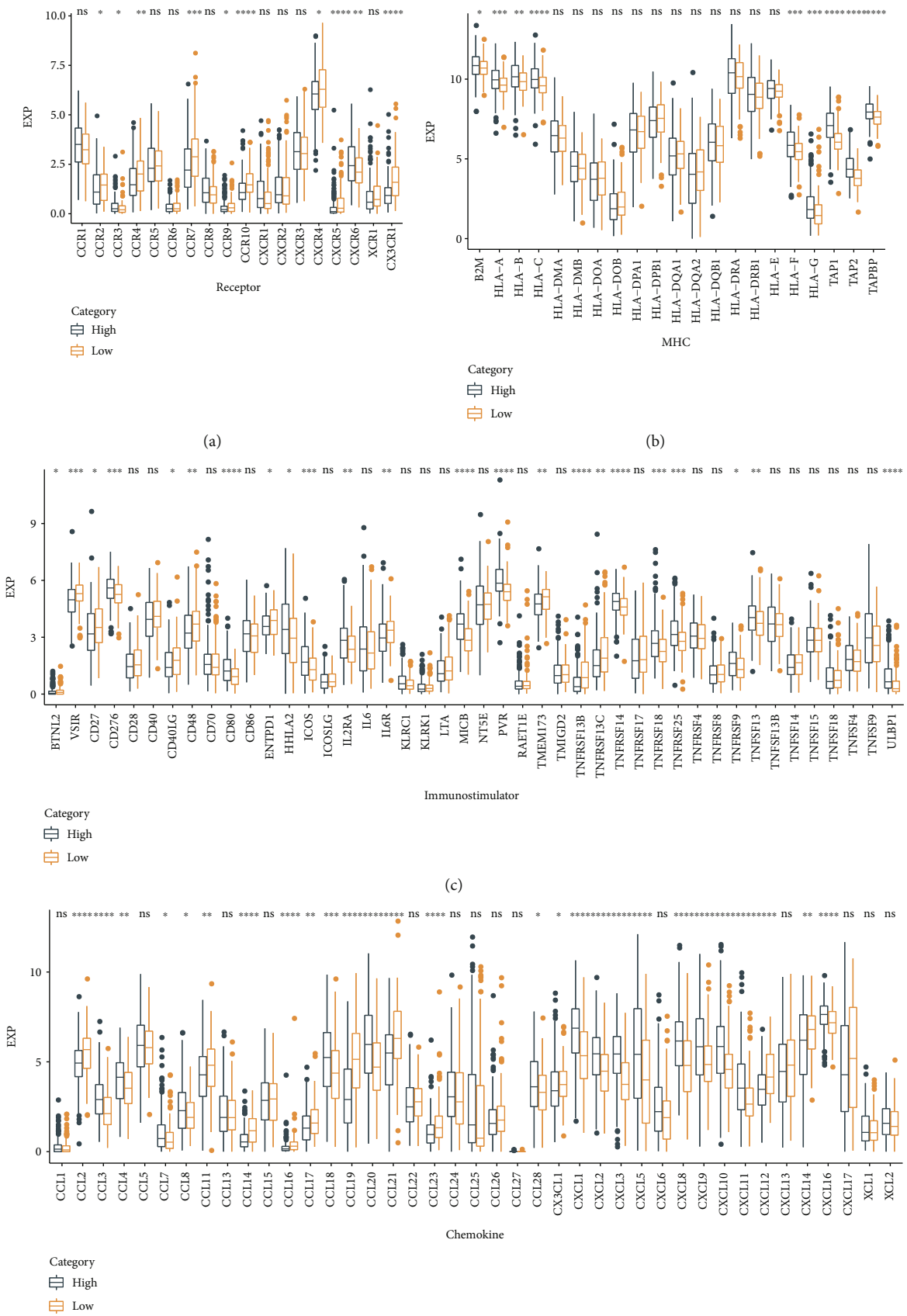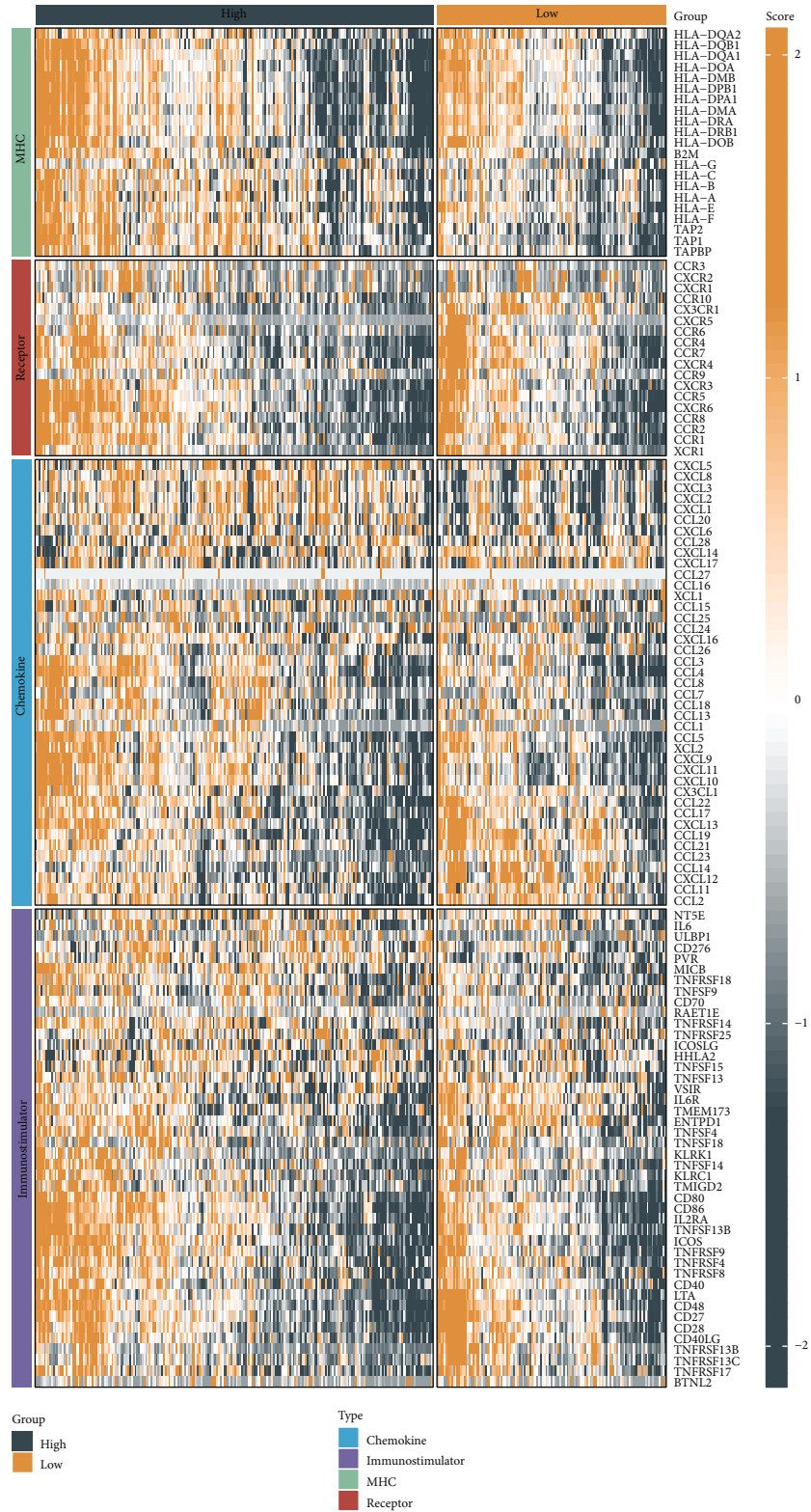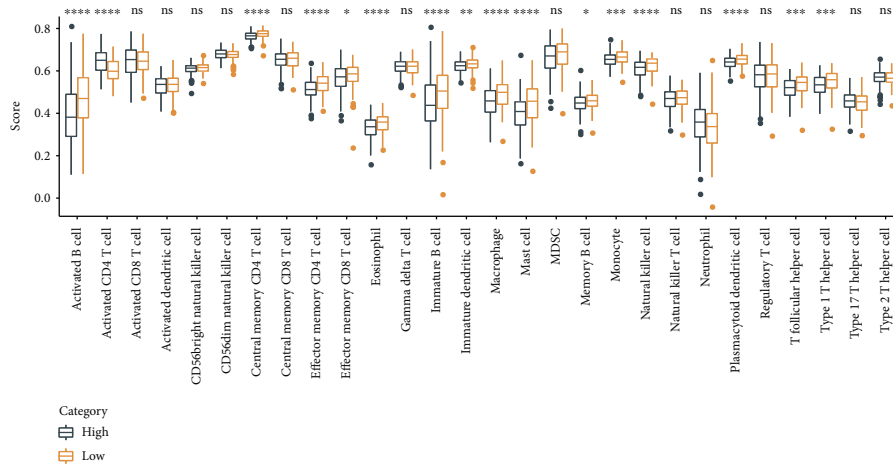
(a)

(b)
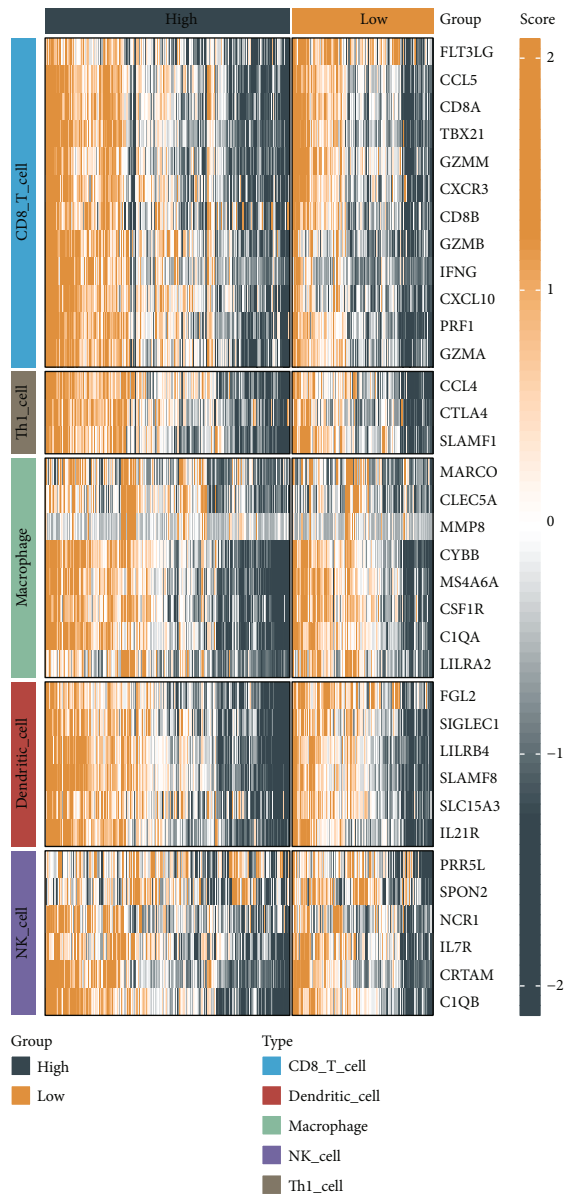

(c)

(d)

FIGURE 4: Continued.

(e)

FIGURE 4: Comparative analysis for immune profiles of NCAPG subgroups in STAD. (a–d) Expression differences of different types of genes (chemokines, receptors, MHC, and immunostimulants) between NCAPG subgroups in STAD; (e) heat map of differential expression among gene types (chemokines, receptors, MHC, and immunostimulants) in STAD (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, $P > 0.05$).
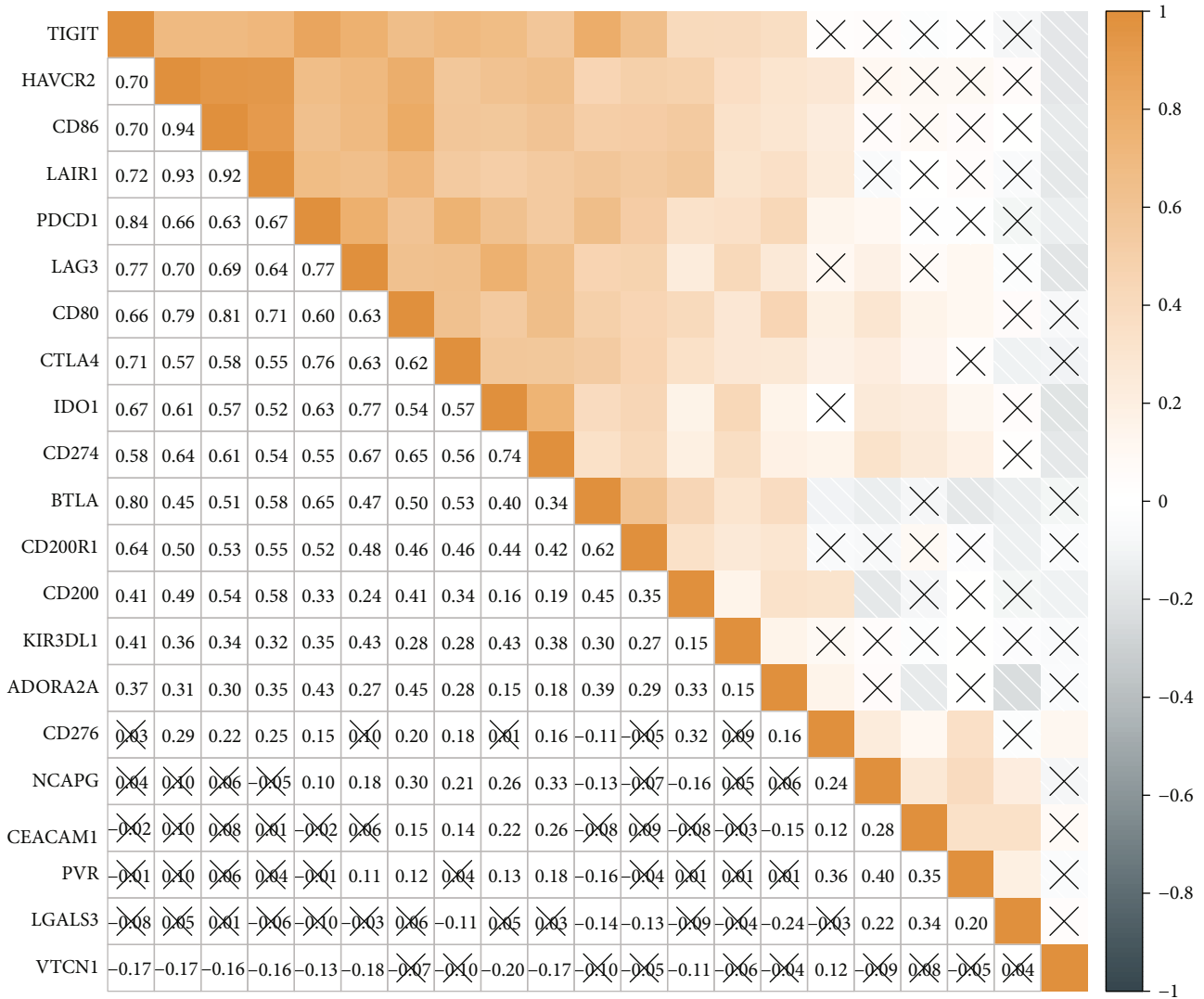
(a)

(b)

Figure 5: Continued.

|	TIGIT	HAVCR2	CD86	LAIR1	PDCD1	LAG3	CD80	CTLA4	IDO1	CD274	BTLA	CD200R1	CD200	KIR3DL1	ADORA2A	CD276	NCAPG	CEACAM1	PVR	LGALS3
HAVCR2	0.70																			
CD86	0.70	0.94																		
LAIR1	0.72	0.93	0.92																	
PDCD1	0.84	0.66	0.63	0.67																
LAG3	0.77	0.70	0.69	0.64	0.77															
CD80	0.66	0.79	0.81	0.71	0.60	0.63														
CTLA4	0.71	0.57	0.58	0.55	0.76	0.63	0.62													
IDO1	0.67	0.61	0.57	0.52	0.63	0.77	0.54	0.57												
CD274	0.58	0.64	0.61	0.54	0.55	0.67	0.65	0.56	0.74											
BTLA	0.80	0.45	0.51	0.58	0.65	0.47	0.50	0.53	0.40	0.34										
CD200R1	0.64	0.50	0.53	0.55	0.52	0.48	0.46	0.46	0.44	0.42	0.62									
CD200	0.41	0.49	0.54	0.58	0.33	0.24	0.41	0.34	0.16	0.19	0.45	0.35								
KIR3DL1	0.41	0.36	0.34	0.32	0.35	0.43	0.28	0.28	0.43	0.38	0.30	0.27	0.15							
ADORA2A	0.37	0.31	0.30	0.35	0.43	0.27	0.45	0.28	0.15	0.18	0.39	0.29	0.33	0.15						
CD276	0.03✕	0.29	0.22	0.25	0.15	0.10✕	0.20	0.18	0.01✕	0.16	−0.11	−0.05✕	0.32	0.09✕	0.16					
NCAPG	0.04✕	0.10✕	0.06✕	−0.05✕	0.10	0.18	0.30	0.21	0.26	0.33	−0.13	−0.07✕	−0.16	0.05✕	0.06✕	0.24				
CEACAM1	−0.02✕	0.10✕	0.08✕	0.01✕	−0.02	0.06✕	0.15	0.14	0.22	0.26	−0.08	0.09✕	−0.08✕	−0.03	−0.15	0.12	0.28			
PVR	−0.01✕	0.10✕	0.06✕	0.04✕	−0.01	0.11	0.12	0.04✕	0.13	0.18	−0.16	−0.04✕	0.01✕	0.01✕	0.01✕	0.36	0.40	0.35		
LGALS3	−0.08✕	0.05✕	0.01✕	−0.06✕	−0.10✕	−0.03✕	0.06✕	−0.11	0.05✕	0.03✕	−0.14	−0.13	−0.09✕	−0.04✕	−0.24	−0.03✕	0.22	0.34	0.20	
VTCN1	−0.17	−0.17	−0.16	−0.16	−0.13	−0.18	−0.07✕	−0.10✕	−0.20	−0.17	−0.10✕	−0.05✕	−0.11	−0.06✕	−0.04✕	0.12	−0.09✕	0.08✕	−0.05✕	0.04✕

(c)

FIGURE 5: Differences in the distribution of immune cell scores between NCAPG subgroups. (a) Differences in immune cell scores between the high and low NCAPG groups. (b) Differences in five effector genes among five TIICs (CD8+ T cells, NK cells, macrophages, Th1 cells, and dendritic cells) between the high and low NCAPG groups. (c) Correlation between NCAPG and immune checkpoints. Colors and values indicate Spearman correlation coefficients (X represents $P > 0.05$) (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, $P > 0.05$).

frequencies of TP53, TTN, SYNB1, RB1, KMT2D, and ARID1A were higher in the low expression group (Figure 3(b)). We then assessed the differences in the TMB of NCAPG between the groups. TMB was markedly higher in the NCAPG high expression group relative to the NCAPG low expression group (Figure 3(c)).

NCAPG expression in the presence of CNV amplifications versus deletions was significantly different from that in the case of the normal copy number, i.e., significantly upregulated expression was associated with CNV amplification while significantly downregulated expression was associated with CNV deficiency ($P < 0.0001$, Figure 3(d)). Meanwhile, the expression of NCAPG correlated negatively with the methylation profile of NCAPG (Figure 3(e), $R = -0.215$).

3.3. Comparative Analysis of the Immune Profiles between NCAPG Expression Groups in STAD. We analyzed the differential expression of four types of genes including chemokines, immunostimulators, MHCs, and receptors, between the two NCAPG expression groups. Among the receptors, genes including CCR7, CCR10, and CXCR5 were significantly higher in the NCAPG low expression group relative to the NCAPG high expression group (Figure 4(a)). In contrast, among the MHCs, genes including HLA-A, HLA-B, and HLA-C were significantly lower in the NCAPG low expression group relative to the NCAPG high expression group (Figure 4(b)). Among the immunostimulators, genes including CD276, CD80, and ICOS were significantly lower in the NCAPG low expression group relative to the NCAPG

(a)

(b)
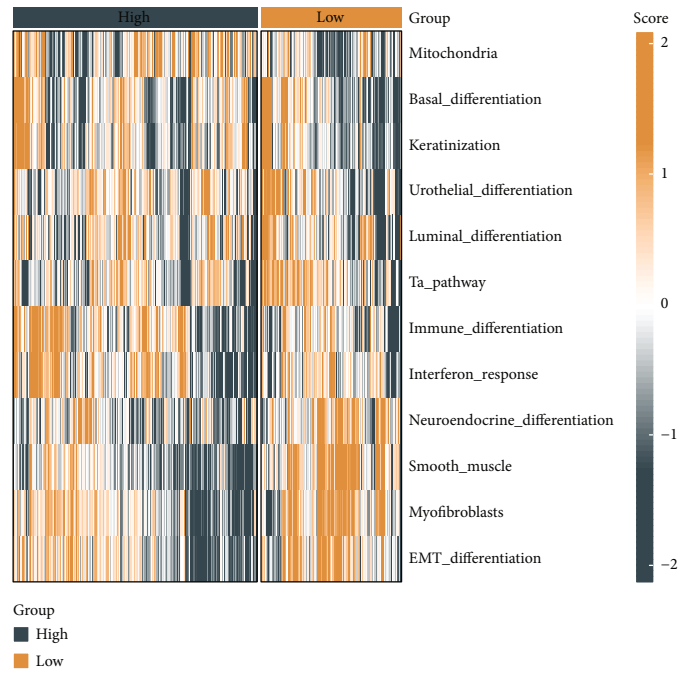
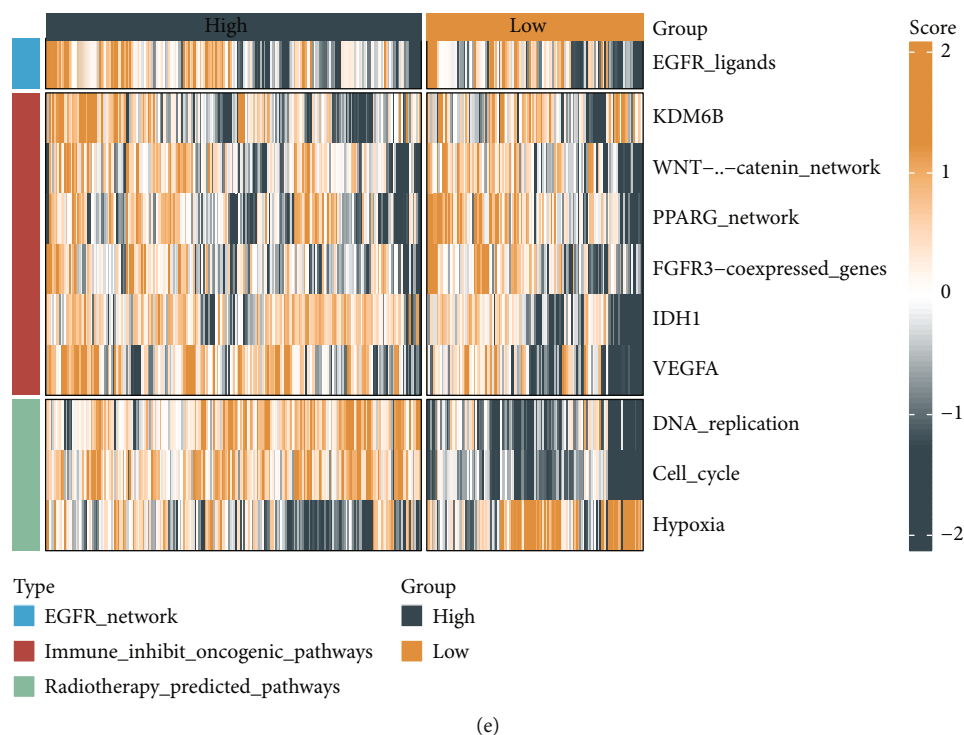FIGURE 6: Continued.

(c)
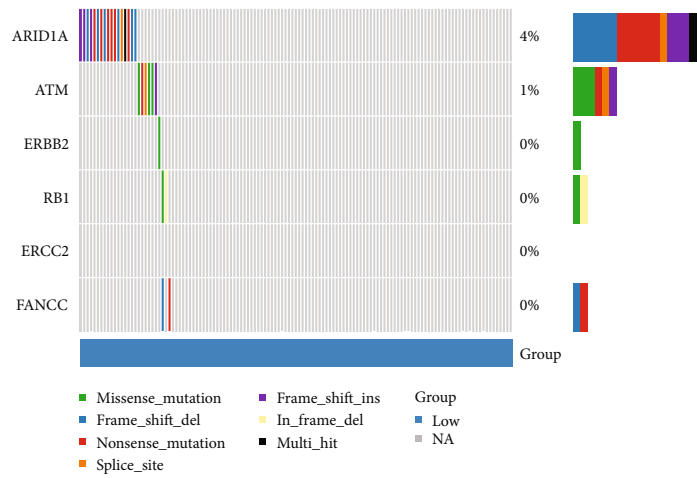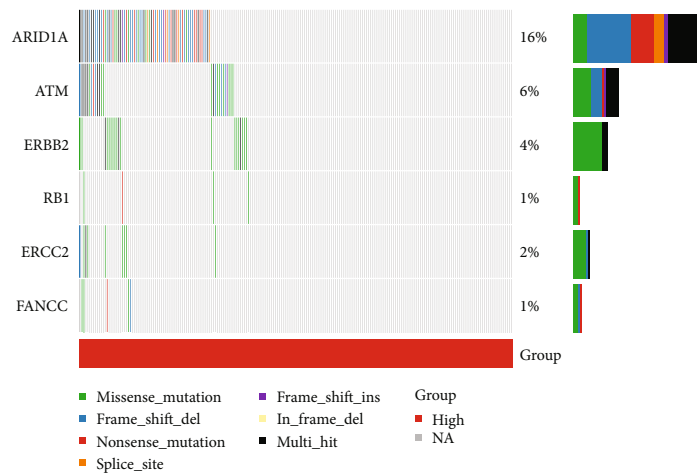
(d)

Figure 6: Continued.

(e)

FIGURE 6: NCAPG immune signature analysis. (a) Correlation between NCAPG and pan-cancer T cell inflammation scores. (b) Pearson correlation analysis between NCAPG and immune checkpoints (X represents $P > 0.05$). (c) Correlation between NCAPG and molecular subtypes and features of STAD using seven different algorithms. (d) Mutational profiles of genes associated with neoadjuvant chemotherapy between low and high NCAPG groups. (e) Correlation between NCAPG and enrichment scores for several treatment features including targeted and radiation therapies (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$).

high expression group (Figure 4(c)). Finally, among the chemokines, CCL3, CCL4, and CCL18 were significantly higher in the NCAPG high expression group than in the low expression group (Figure 4(d)). Overall, the four key immune checkpoint genes differed significantly between the two evaluated groups (Figure 4(e)).

We compared the distributional differences of 28 immune cell scores between the NCAPG group, and 16 of them showed significant differences with high scores in the NCAPG low expression group (Figure 5(a)). We then performed immune infiltration analysis in STAD, calculated the association between immune infiltration scores and NCAPG expression, and performed a differential analysis of marker genes in five cell types, namely, dendritic_cell, CD8_T_cell, Th1_cell, NK_cell, and macrophage. In the NCAPG low expression group, FLT3LG was markedly upregulated (Figure 5(b)). We also assessed the association between immune checkpoints and NCAPG expression; significant slight positive correlations with immune checkpoints including CD80, CTLA4, IDO1, and CD274 were recorded (Figure 5(c)).

3.4. NCAPG Expression Predicts the Clinical Responses to Immune Checkpoint Blockade (ICB) in STAD. The correlation between NCAPG expression and pan-cancer T cell inflammation score was assessed; but no significant correla-

tion was observed (Figure 6(a)). Additionally, in subgroups with different ICB responses, the correlation between NCAPG expression and immune features (immune checkpoints, immunotherapy-related features, and expression of TIIC effector genes and immunomodulators) (Figure 6(b)) was assessed; NCAPG expression was not correlated with any of these features.

The association of scores of STAD tumors with immune-related pathways was evaluated; between the two NCAPG expression groups, significant differences in the associated immune pathways were noted (Figure 6(c)). For instance, smooth muscle and myofibroblasts pathways showed a predominant negative association with high NCAPG expression, while the trend was opposite in the low NCAPG group. We compared the mutation frequencies in ARID1A, RB1, ERBB2, and FANCC, which are associated with radiotherapy, between the NCAPG groups (Figure 6(d)); ARID1A, ATM, ERBB2, RB1, ERCC2, and FANCC were more frequently mutated in the high NCAPG group. We compared the differences in EGFR_network, immune_inhibit_oncogenic_pathways, and radiotherapy_predicted_pathways between the groups (Figure 6(e)). The high NCAPG group correlated positively with DNA_replication and cell_cycle pathways among the radiotherapy_predicted_pathways type, while the trend was opposite in the low NCAPG group.
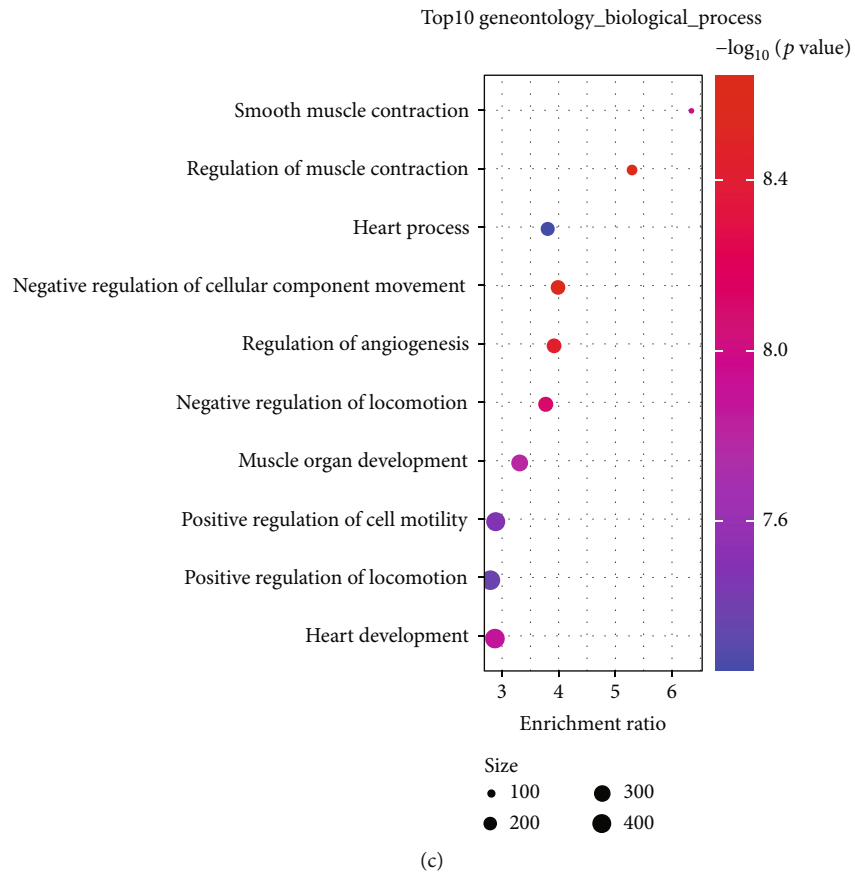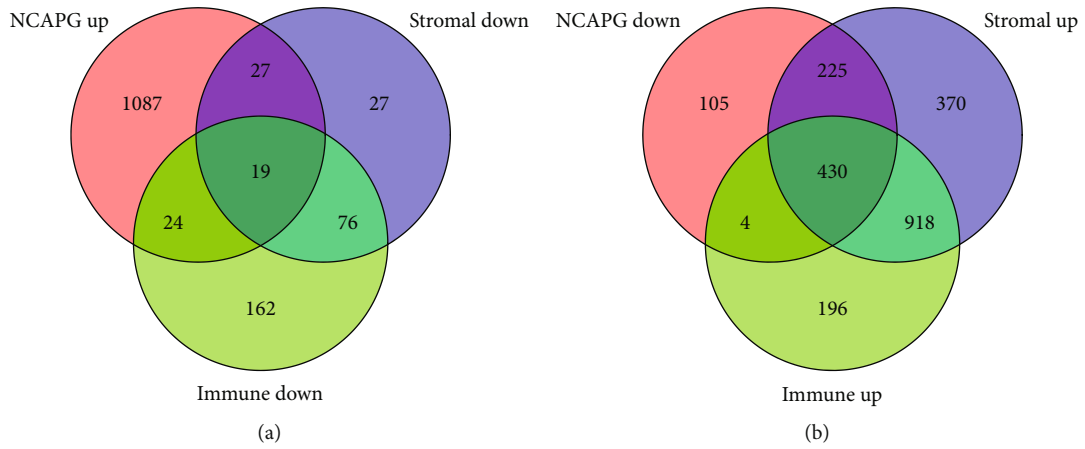
(a)

(b)

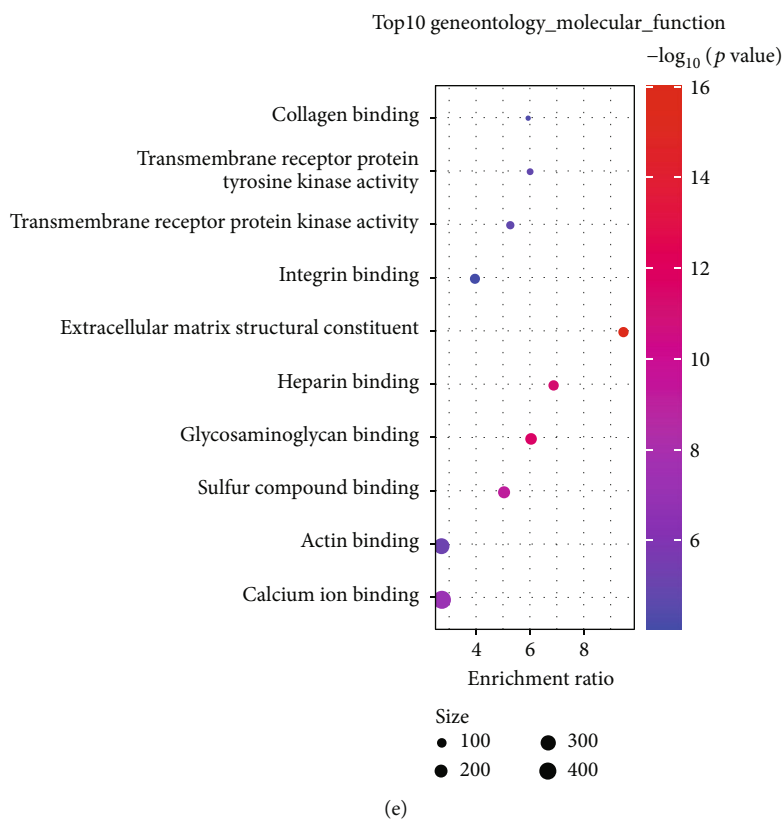Top10 geneontology_biological_process

(c)

Figure 7: Continued.

Top10 geneontology_cellular_component

(d)

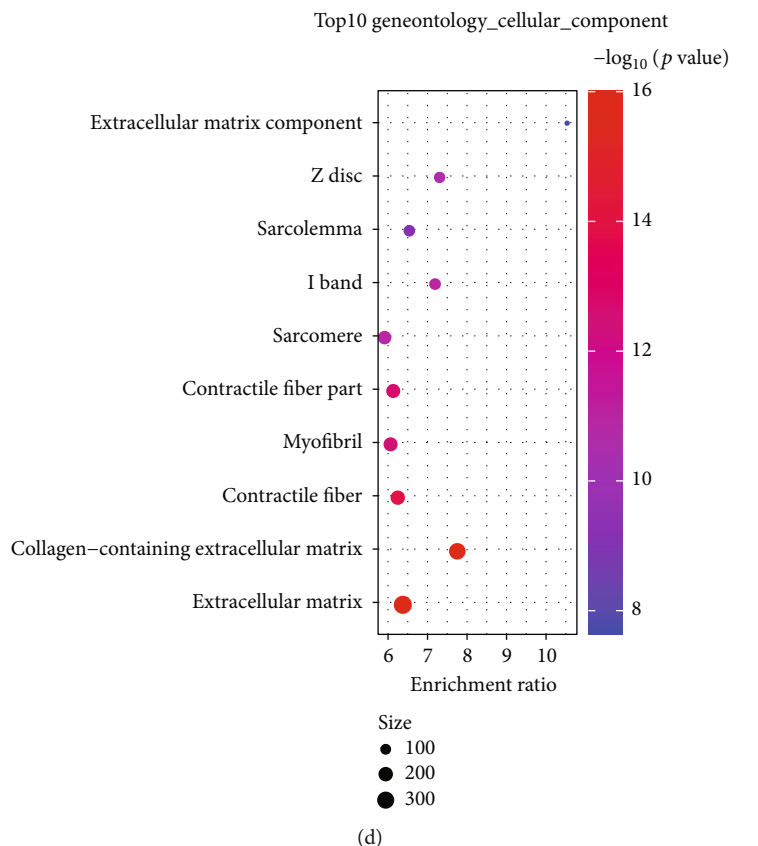Top10 geneontology_molecular_function

(e)

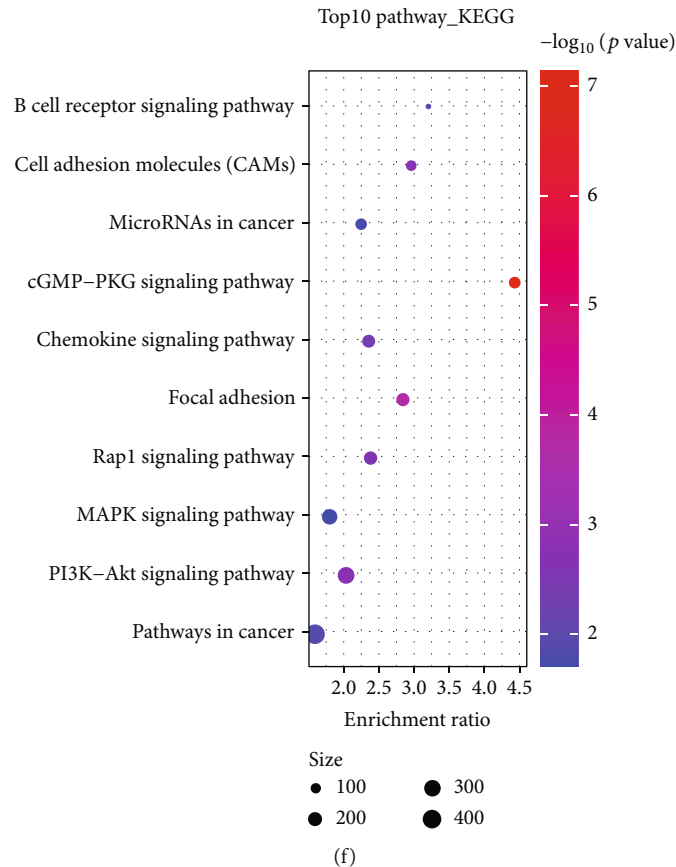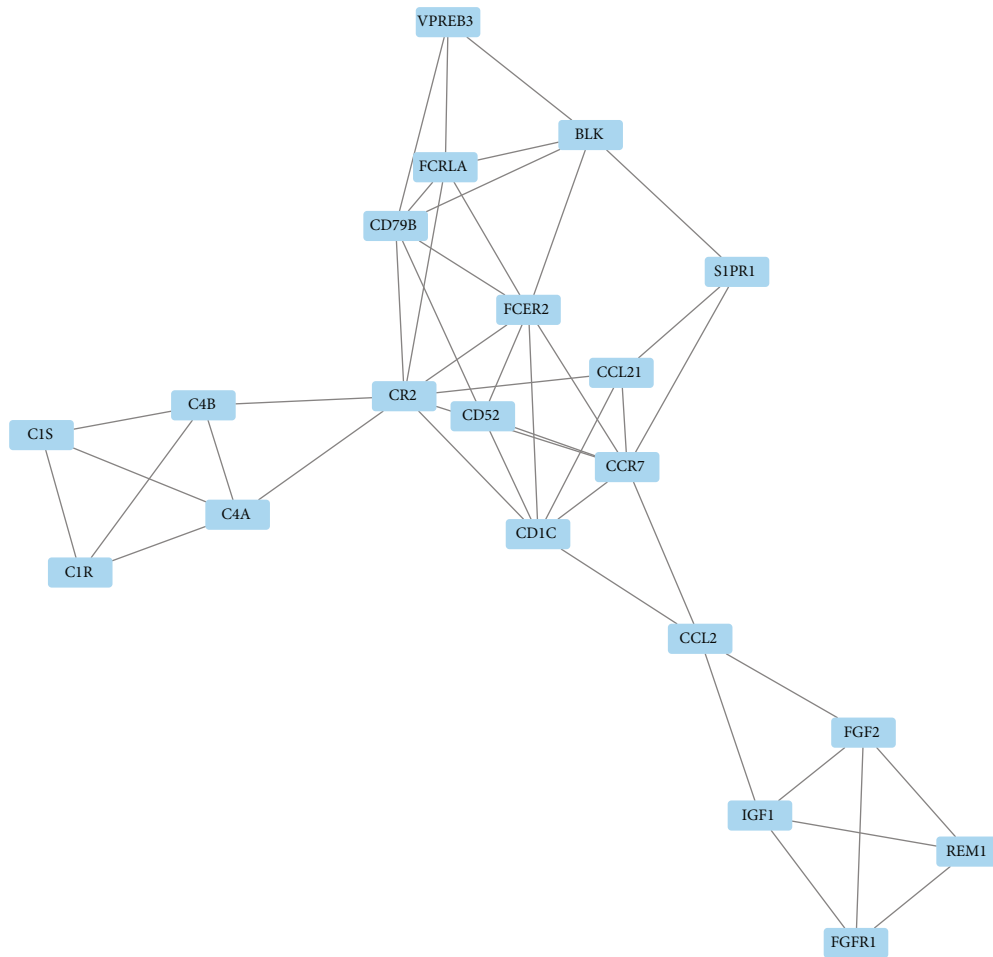FIGURE 7: Continued.

Top10 pathway_KEGG

(f)

FIGURE 7: Immune-related differential gene expression analysis. (a) Intersection of upregulated genes among NCAPG, StromalScore, and ImmuneScore groups; (b) the intersection of downregulated genes among NCAPG, StromalScore, and ImmuneScore groups; (c–f) GO annotation and KEGG functional enrichment analyses for DEGsin NCAPG, StromalScore, and ImmuneScore groups.

3.5. Identification of Immune-Related DEGs and PPI Network Analysis. The ESTIMATE algorithm was used to compute the immune and stromal scores, and two groups corresponding to each of them were classified. In the correlation analysis, expectably, we observed a significantly negative correlation of NCAPG expression with immune score and stromal score ($R = -0.178$ and -0.352, respectively, Fig. S2). We next divided STAD samples into two groups according to the cutoff values determined by survminer R package of NCAPG expression, immune score, and stromal score, respectively. Differential expression analysis identified the upregulated and downregulated DEGs between low and high NCAPG expression groups, low- and high-immune score groups, and low- and high-stromal score groups. Given that a negative correlation of NCAPG expression with immune score and stromal score was observed, we therefore performed an intersection analysis among upregulated DEGs of NCAPG group and downregulated DEGs of immune and stromal groups. A total of 19 DEGs were overlapped in the above three groups (Figure 7(a)). By using the same analysis, we identified 430 overlapping in downregulated DEGs of NCAPG group and upregulated DEGs of immune and stromal groups (Figure 7(b)).

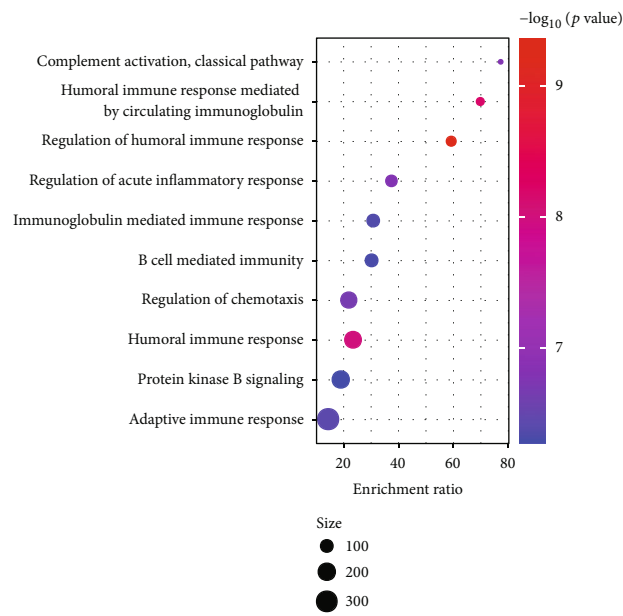GO annotation and KEGG functional enrichment analyses for the DEGs were performed using WebGestaltR. Genes were closely related to pathways of tumorigenesis and immune responses (Figures 7(c)–7(f)), including regulation of angiogenesis, B cell receptor signaling pathway, microRNAs in cancer, and pathways in cancer. STRING was used for PPI analysis of 449 obtained DEGs. Cytoscape was utilized for network visualization and the MCODE plugin for identifying the key clusters. Five clusters with more than 10 genes (Mcode1, 2, 3, 5, and 9) were identified. KEGG functional enrichment analysis using WebGestaltR was performed to identify their functions. The Mcode3 module was closely related to immune and tumor pathways, including breast cancer, proteoglycans in cancer, and chemokine signaling (Figure 8).

3.6. IRS Model for STAD. After the above analysis, 449 immune-related differential genes were obtained, and univariate analysis revealed 248 prognosis-related genes ($P < 0.05$). The top genes were further selected using LASSO, and 10 genes were obtained according to the minimum lambda = 0.07841627 (Figure 9(a)). These were subjected to multifactorial analysis, and their risk coefficients were obtained (Figure 9(b)). IRS risk models were constructed according to the expression of these 10 genes. We calculated the risk score for each sample in the training and validation datasets of GSE66229 and divided the samples into high and

(a)

Top10 geneontology_biological_process

(b)

FIGURE 8: Continued.

All geneontology_cellular_component
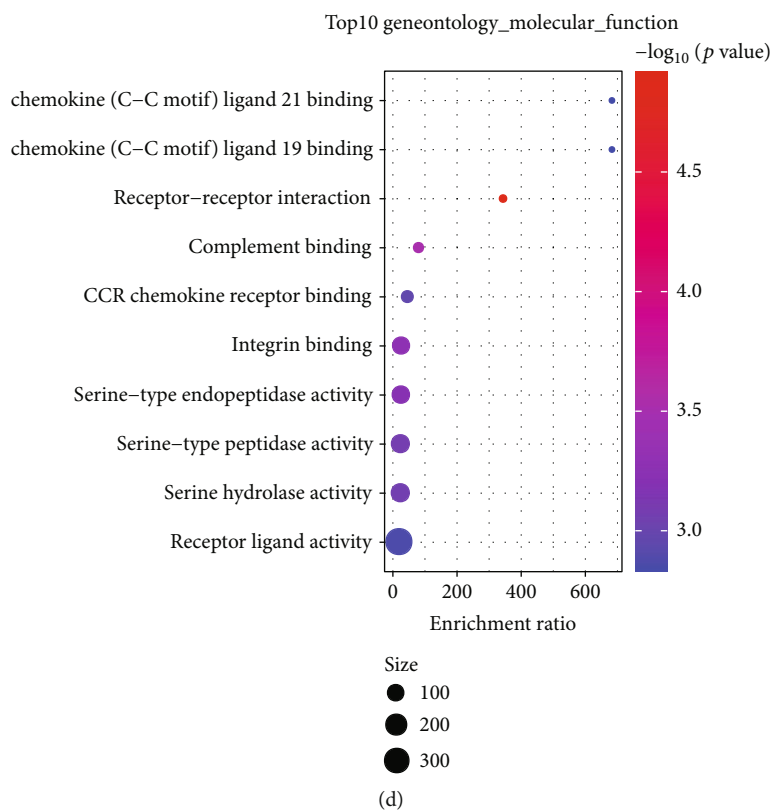

(c)

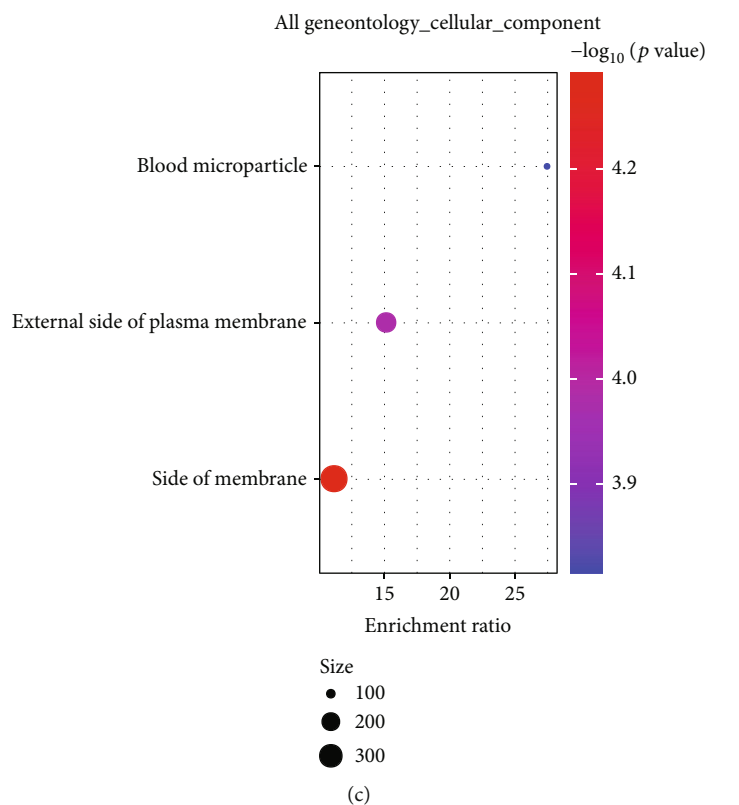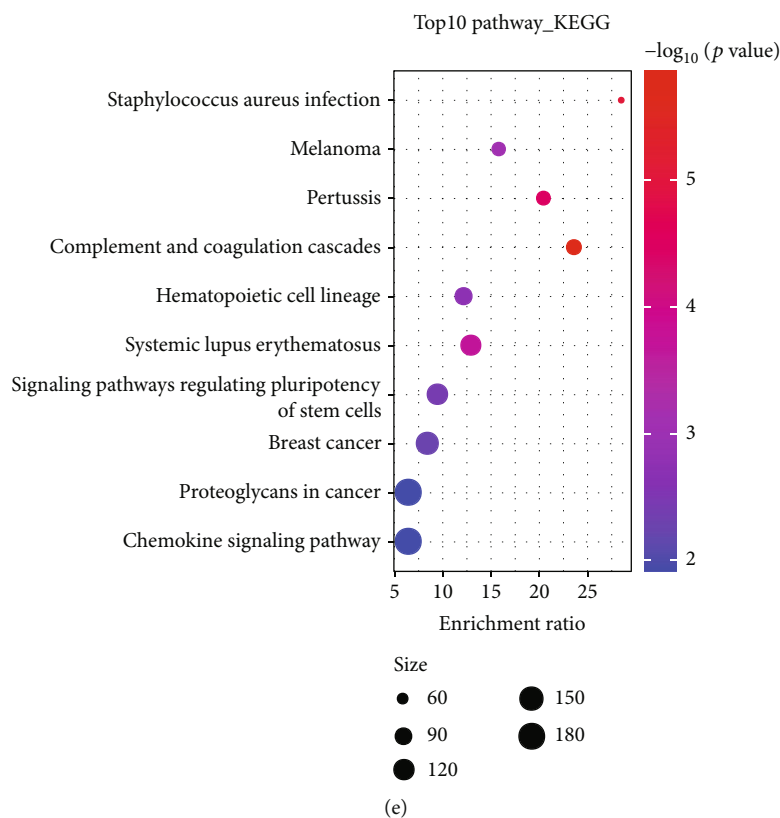Top10 geneontology_molecular_function

(d)

FIGURE 8: Continued.

FIGURE 8: Module and functional analyses. (a) PPI network analysis graph for module Mcode3; (b–e) GO and KEGG functional enrichment analyses for genes in module Mcode3.

low-risk groups according to the best cutoff. KM and ROC curves were plotted, wherein the AUC of one-year survival was 0.75, 0.81 for three years, and 0.83 for five years in the training set, while the corresponding values of one, three, and five years were 0.62, 0.68, and 0.67 (Figures 9(c) and 9(d)) in the validation set. Both in training and validation sets, the survival rate was higher for the low-risk group ($P < 0.0001$). The IRS model was validated using the entire GSE66229 and TCGA cohorts (Figures 9(e) and 9(f)). The STAD tumor samples could be clearly classified into two risk groups with different prognoses ($P < 0.0001$); the low-risk group showed a higher survival rate. For GSE84437, GSE26942, GSE13861, GSE28541, and GSE26253 datasets, the IRS scores for each sample were recalculated and KM and ROC curves were plotted. The corresponding AUCs for one-year survival were 0.64, 0.63, 0.66, 0.71, and 0.65; 0.65, 0.65, 0.8, 0.82, and 0.68 for three years, and 0.65, 0.66, 0.76, 0.89, and 0.69, for five years, respectively (Fig. S3).

3.7. The Relation between Risk Score and Tumor Stemness. Tumor stemness has been revealed to be correlated with prognosis, which is linked with tumor microenvironment and oncogenic pathways such as Wnt/β-catenin and EMT pathways [23]. We calculated the mRNAsi scores of STAD samples in different datasets and compared their distribution in high- and low-risk groups. The results showed that the low-risk group had significantly higher mRNAsi scores

than the high-risk group in all independent datasets including GSE66229, TCGA, GSE84437, GSE26942, GSE13861, GSE28541, and GSE26253 datasets (Figure 10(a)). In addition, we observed that NCAPG expression was positively correlated with mRNAsi ($R = 0.728$, $P < 0.0001$, Figure 10(b)). The results suggested that high degree of tumor stemness was associated with favorable prognosis in STAD.

3.8. Comparative Performances of IRS and TIDE. To predict, evaluate, and compare efficacy scores for immunotherapy, the IMvigor210, GSE91061, GSE78220, and GSE135222 cohorts were selected as these comprised patients who underwent immunotherapy. We calculated IRS and TIDE score for each tumor sample in these cohorts and compared the efficiency of IRS and TIDE in the predicting the prognosis of the patients receiving immunotherapy. The samples were categorized into high- and low-scoring groups according to the best cutoff of IRS and TIDE scores. As a result, IRS was robust to distinguish high- and low-risk patients while TIDE score was only effective to predict the prognosis in the IMvigor210 dataset (Figures 11(a)–11(h)). The comparison of AUC values also suggested IRS had a superior performance than TIDE score in classifying the patients treated by immunotherapy into different risk groups (Figures 11(i)–11(l)). Therefore, IRS had a potential to predict the prognosis for not only untreated patients but immunotherapy-treated patients in STAD.
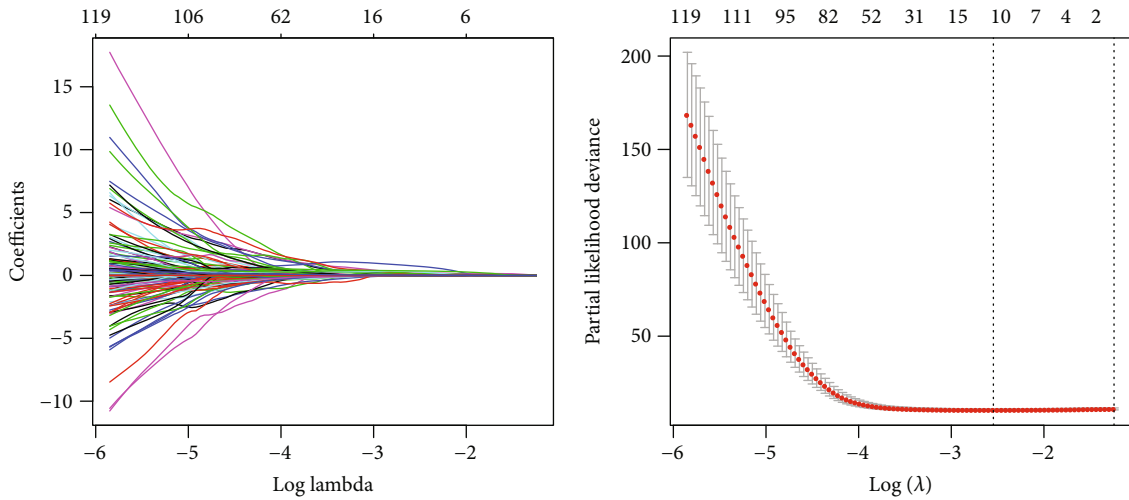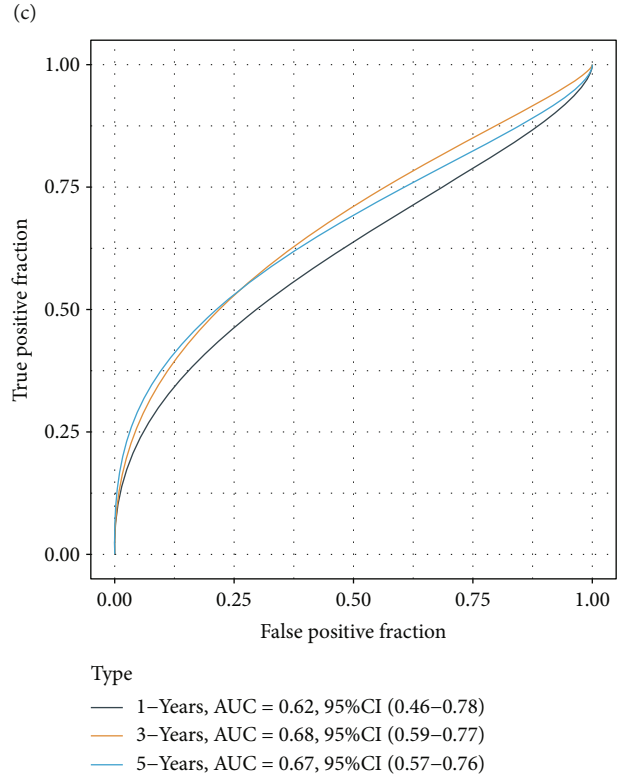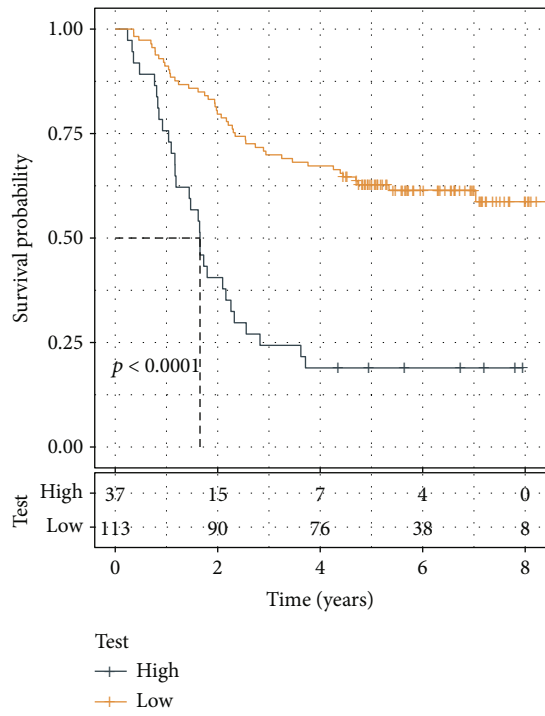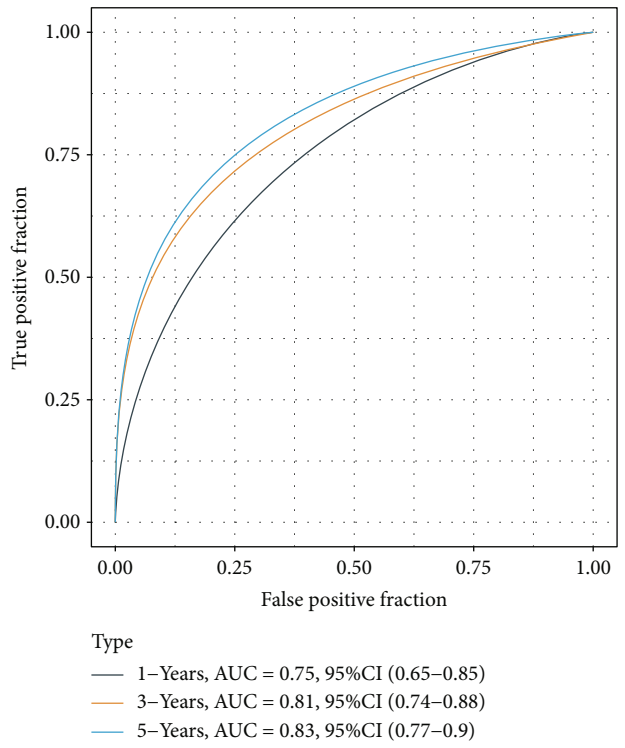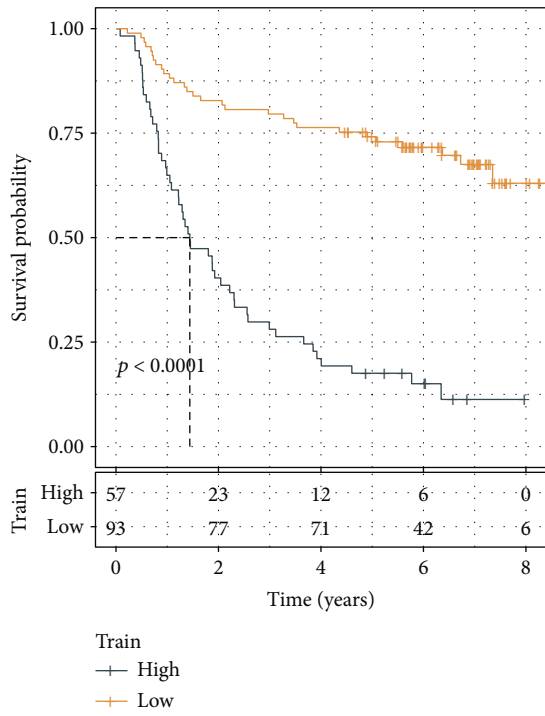
(a)
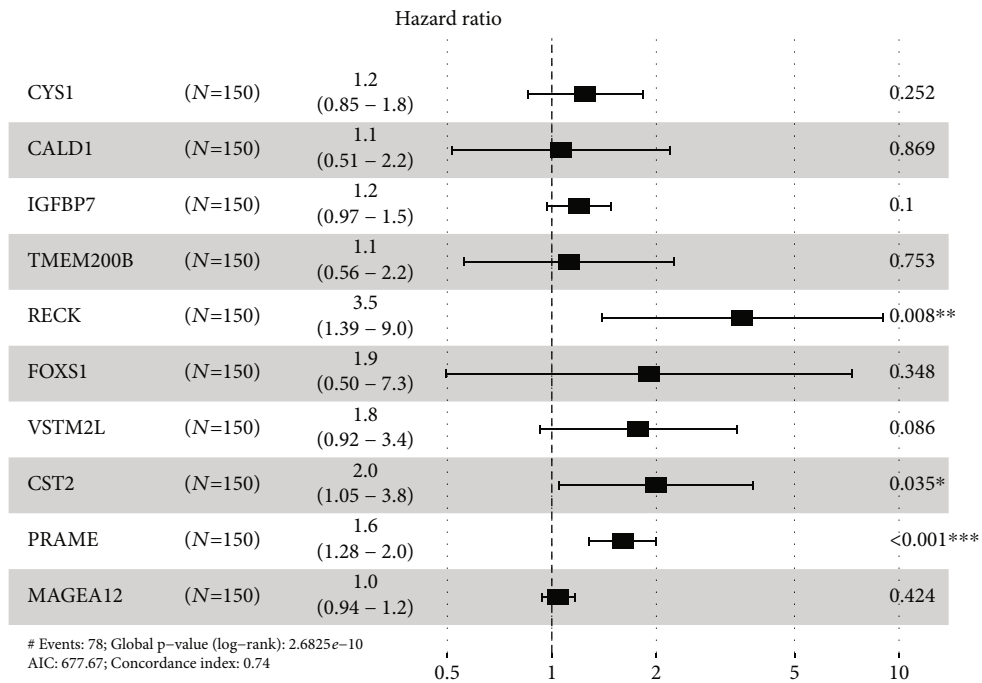
(b)

FIGURE 9: Continued.

(c)
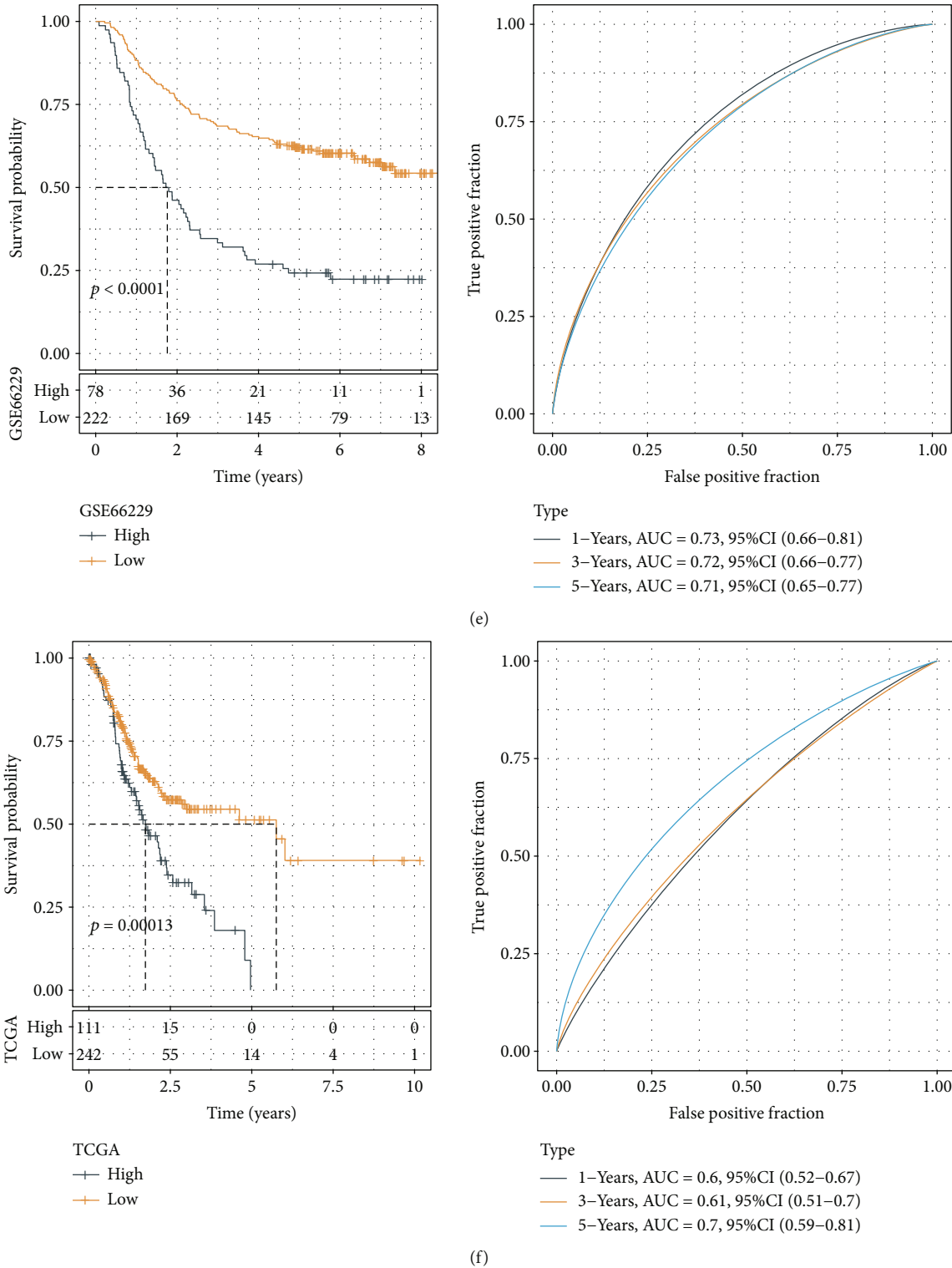

(d)

Figure 9: Continued.

(e)

(f)

FIGURE 9: Construction of IRS for STAD. (a) Distribution of LASSO coefficients for 40 prognostic RNAs in the GEO training cohort. Coefficient profiles were plotted based on log (λ) sequences. (b) Multifactorial results of genes for the final IRS model. (c) KM and ROC analyses for the IRS model using the GEO training dataset. (d) KM and ROC analyses for the IRS model using the GEO validation dataset. (e) KM and ROC analyses for the IRS model using the entire GEO dataset. (f) KM and ROC analyses for the IRS model using the entire TCGA cohort.
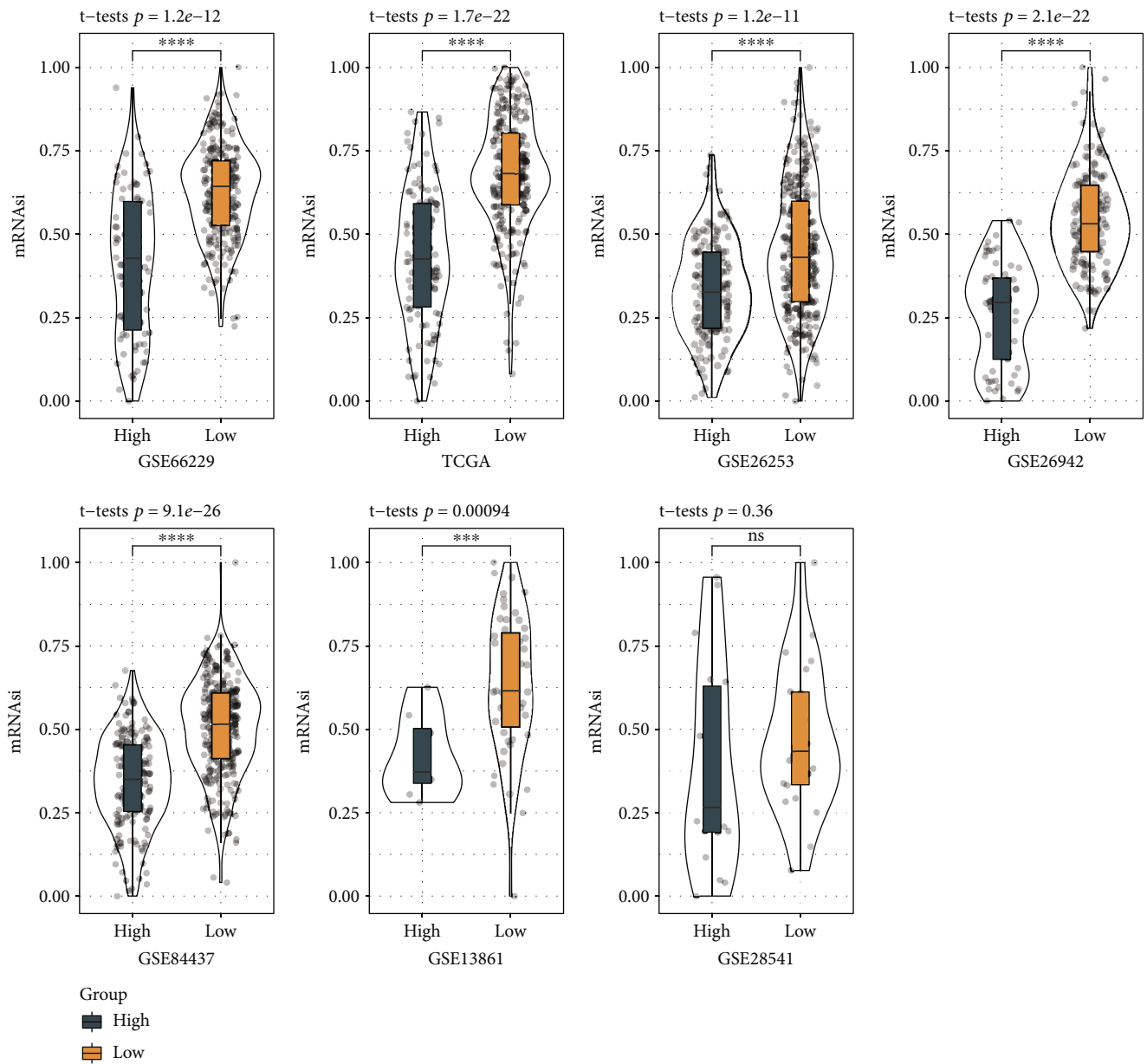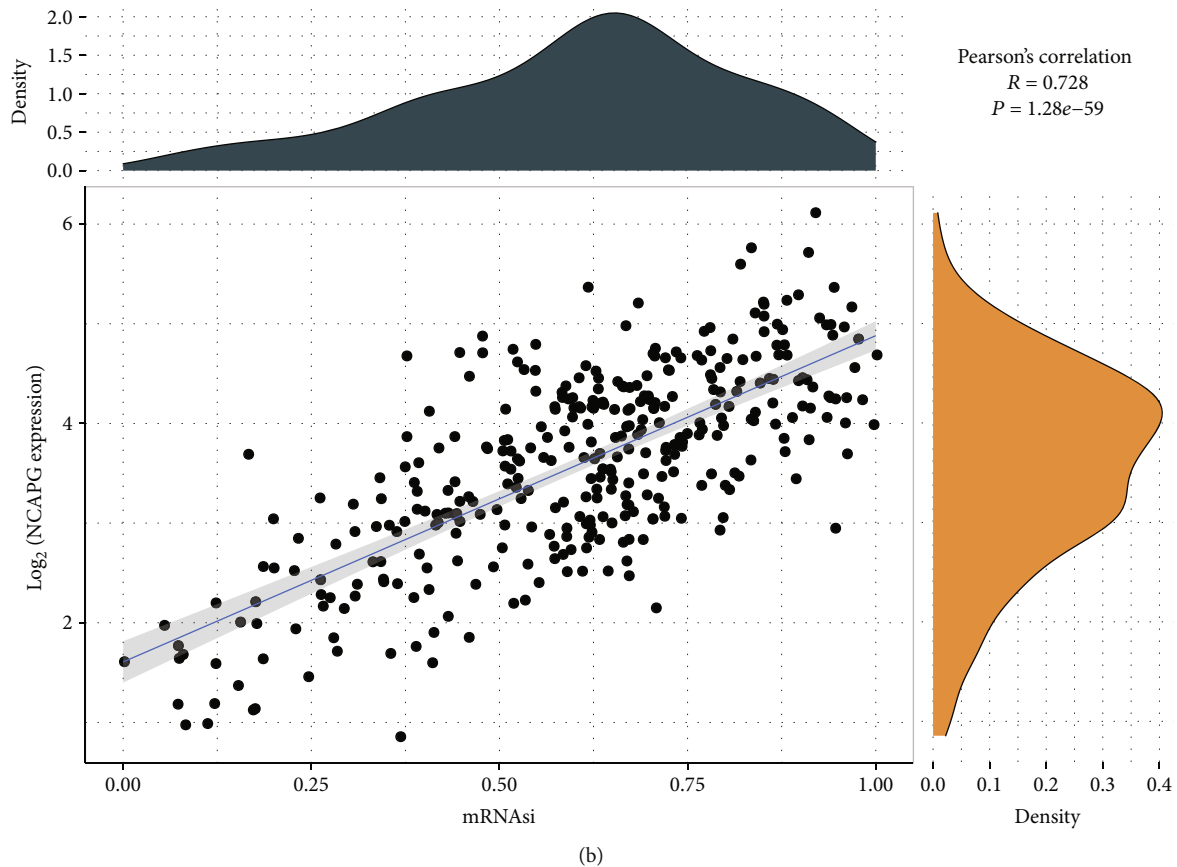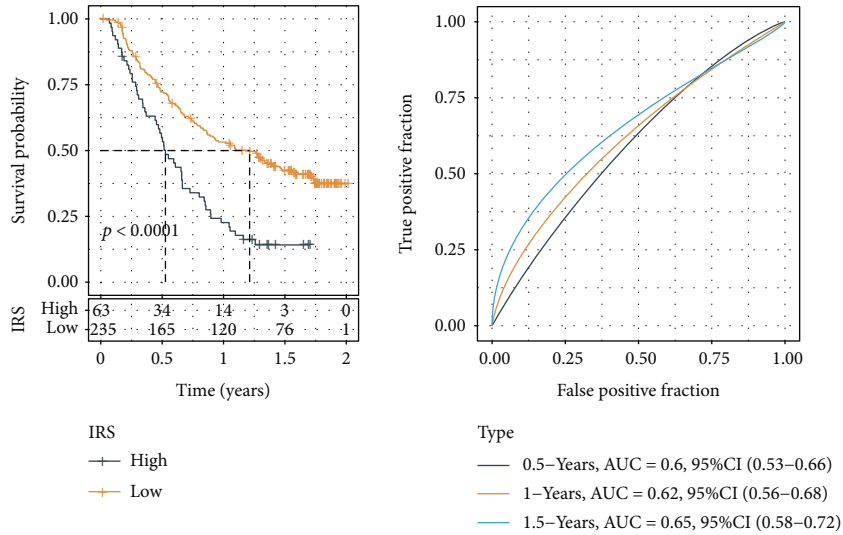
FIGURE 10: Continued.

Figure 10: The relation between risk score and tumor stemness. (a) Comparison of tumor stemness (mRNAsi) between high- and low-risk groups in GSE66229, TCGA, GSE84437, GSE26942, GSE13861, GSE28541, and GSE26253 datasets. Student's t-test was conducted. (b) Pearson correlation analysis between NCAPG expression and mRNAsi. ns: not significant. ***$P < 0.001$ and ****$P < 0.0001$.

4. Discussion

In 21 cancer types, especially, STAD, NCAPG expression was significantly higher in tumor samples relative to the para-cancer samples; it was positively correlated with all four types of immune-related genes. Therefore, we used STAD as the case study. STAD is one of the most prevalent gastrointestinal malignancies, ranking fifth among incidences and third in mortality worldwide [24, 25]. Although 5-year survival rates are 90-97% when diagnosed and treated early, nearly 30% of patients with STAD are diagnosed at the advanced and metastatic stages [24, 25]. The malignant feature of STAD is determined not only by the activation and recruitment of immune and stromal cells, the major components necessary for tumor development and progression, in the tumor-associated microenvironment but also by the intrinsic activity of cancer cells [26]. Growing evidence confirms that the tumor microenvironment exerts an important effect on the development, progression, prognosis, and immune therapeutic responses in STAD [27, 28]. Stromal cell functions and interactions with tumor cells contribute to the progression, invasion, and spread of the tumor cells [29]. Stromal cells can secrete growth factors, cytokines, and chemokines which have a significant impact on tumor characteristics [30]. In STAD, NCAPG was significantly pos-

itively correlated with some chemokines. Therefore, NCAPG is a potentially new and reliable biomarker to predict the development and prognosis of STAD.

Several recent studies have focused on tumor immunology, and many immune checkpoint inhibitors have been developed which show robust and durable responses in patients with different cancers [31, 32]. This is consistent with the results reported herein and corroborates the reliability of our findings. In STAD, immune checkpoint genes including CD80, CTLA4, IDO1, and CD274 correlated positively with NCAPG expression. Based on clinical trials for immune checkpoint inhibitors for various cancers, immune cell infiltration in the tumor microenvironment is now considered valuable information for predicting prognosis and responses to immunotherapy [33, 34]. Therefore, we performed a comprehensive analysis of the overall immune cell infiltration in the tumor microenvironment by estimating the distributional differences of 28 immune cell scores in STAD among NCAPG subgroups. Twenty-one immune cell types showed significant differences between NCAPG grouping; the NCAPG low expression group had high immune scores. Macrophages act as immunosuppressive cells and they impede the activation of NK and CD8+ T cells; most of their marker genes were highly expressed in the NCAPG low expression group [35]. Immunosuppressive

(a)
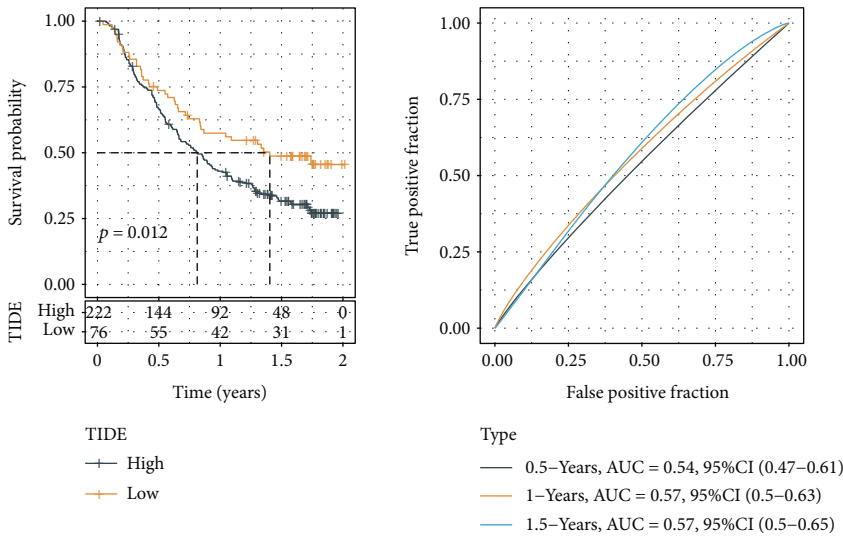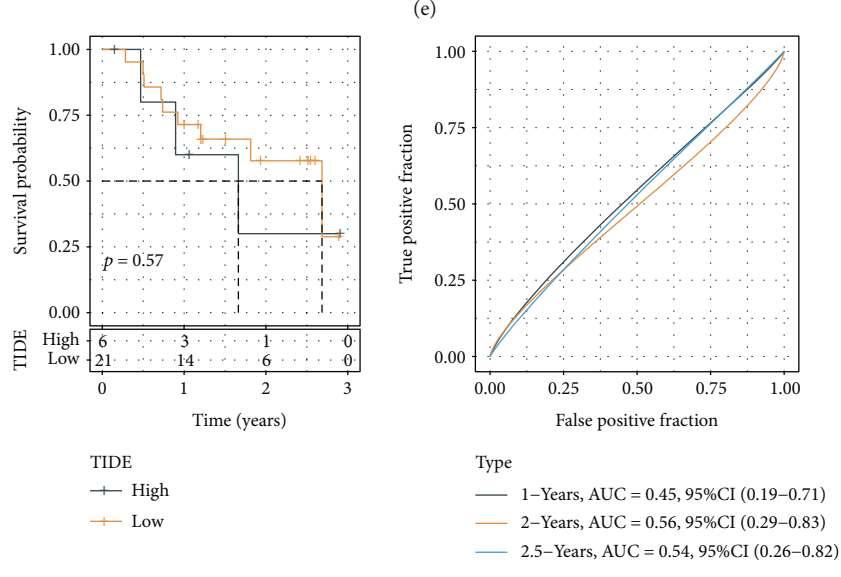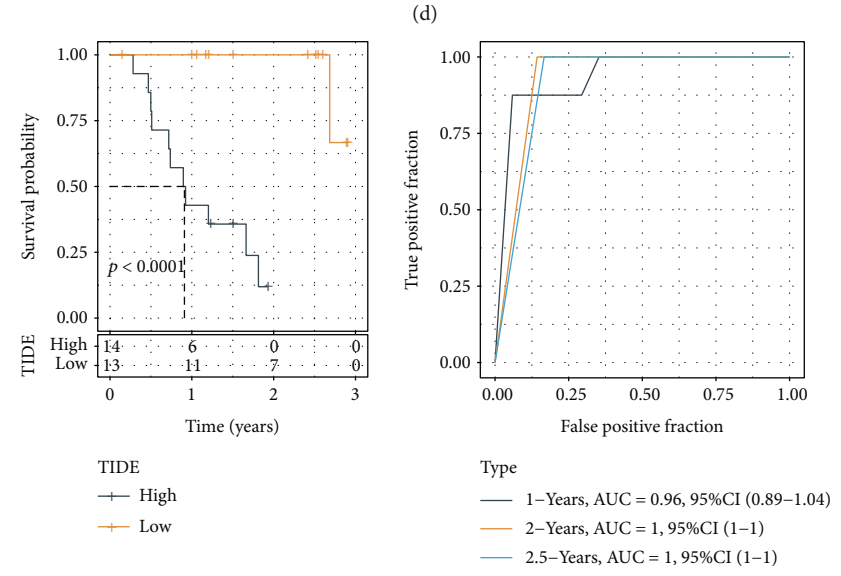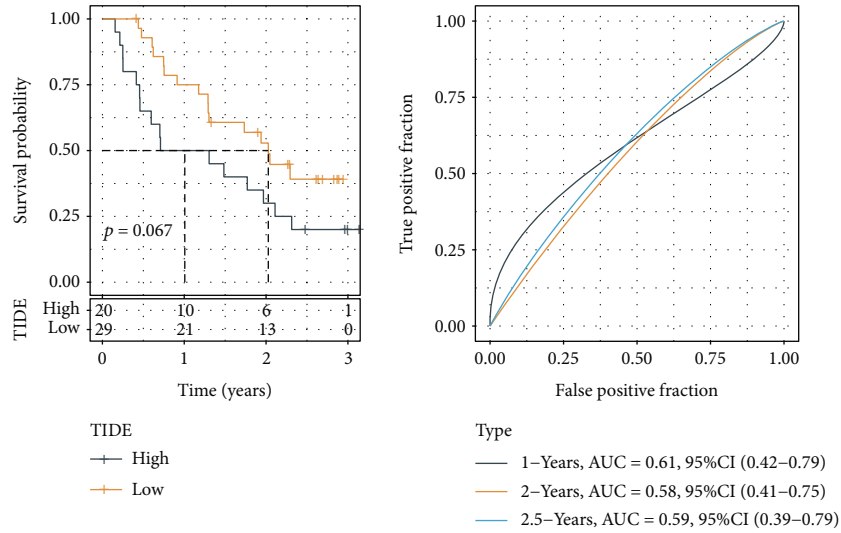

(b)
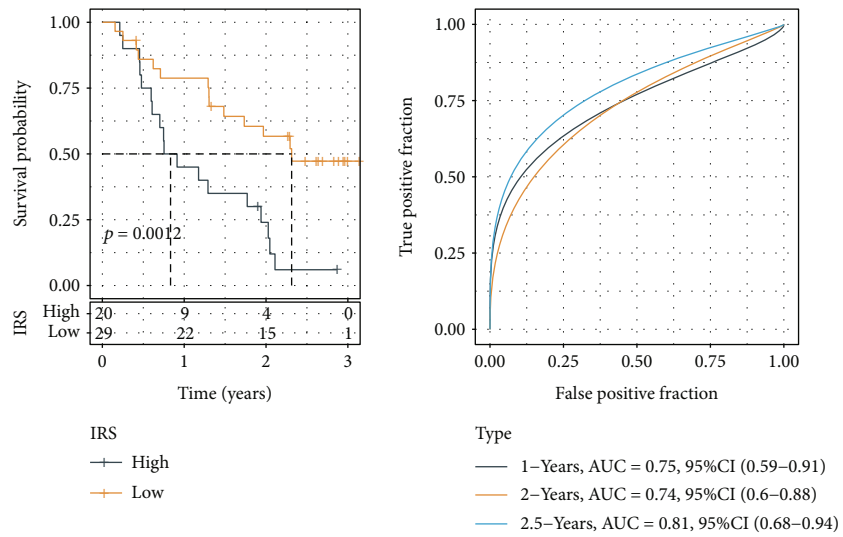
(c)

FIGURE 11: Continued.

(d)

(e)

(f)

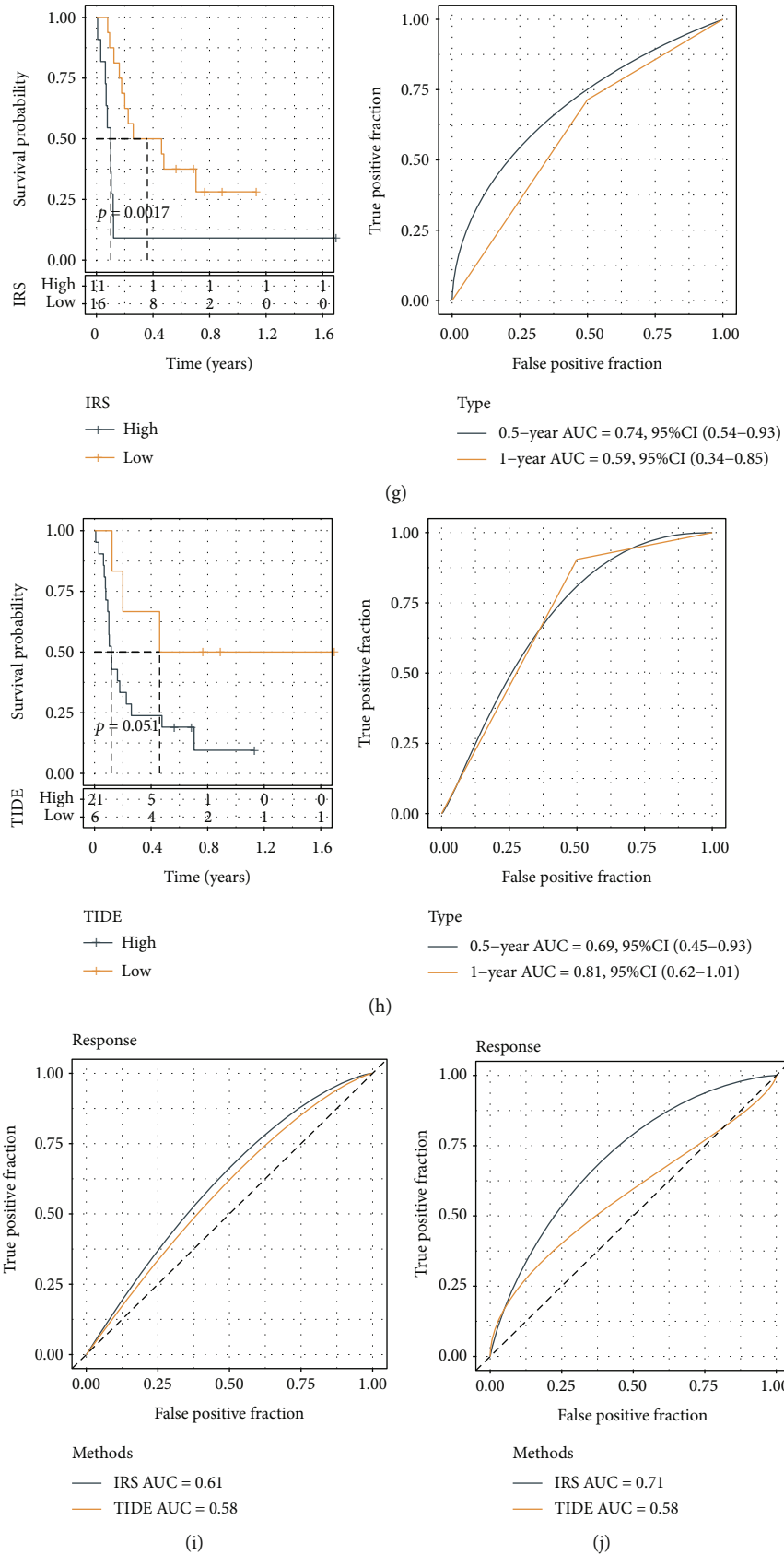FIGURE 11: Continued.
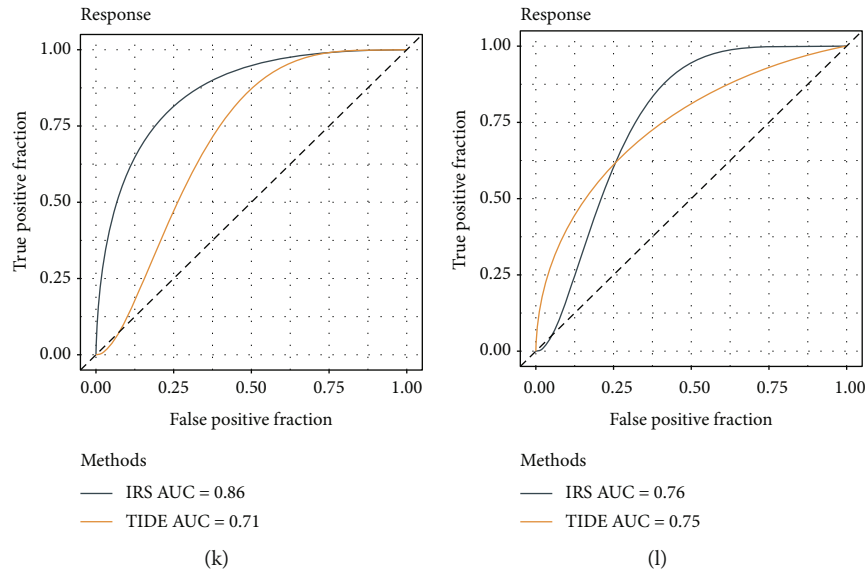
(g)

(h)

(i)

(j)

Figure 11: Continued.

(k)

(l)

FIGURE 11: Comparative analysis of IRS and TIDE. (a) IRS survival and ROC curves for the IMvigor210 dataset. (b) TIDE survival and ROC curves for the IMvigor210 dataset. (c) IRS survival and ROC curves for the GSE91061 dataset. (d) TIDE survival and ROC curves for the GSE91061 dataset. (e) IRS survival and ROC curves for the GSE78220 dataset. (f) TIDE survival and ROC curve in the GSE78220 dataset. (g) IRS survival and ROC curves for the GSE135222 dataset. (h) TIDE survival and ROC curves for the GSE135222 dataset. Comparison of IRS and TIDE ROC curves in (i) IMvigor210, (j) GSE91061, (k) GSE78220, and (l) GSE135222 datasets.

cells can respond to changes in immune cell abundances and occupy a central position in the tumor immune microenvironment. Therefore, we hypothesized that the poor prognosis in low NCAPG expression group may be related to this tumor immunosuppressive microenvironment.

Cancer stem cells (CSCs) compose only a fraction of 0.01-2% of tumor tissues in most tumors, but have great potential in initiating tumor growth [23]. It has been shown that tumor stemness varies in different cancer types [36]. In STAD, mRNAsi is suggested as a protective factor that high mRNAsi is associated with good prognosis [36, 37]. In our study, we found a significantly positive correlation of NCAPG expression with mRNAsi and high NCAPG expression group had a good prognosis, which was consistent with the previous findings. In high- and low-risk groups divided by the immune risk model, mRNAsi was higher in the low-risk group than in the high-risk group, which further supported mRNAsi as a protective factor in STAD. At the same time, the results also revealed the important role of NCAPG in tumor stemness. Experimental work has illustrated that quiescent tissue SCs escape from immune elimination resulting from systemic downregulation of MHC class I and TAP proteins that are responsible for antigen presentation [38]. Our results showed that the low NCAPG expression group had lower expression of some MHC class I genes such as HLA-A, HLA-B, HLA-C, HLA-F, HLA-G, TAP1, and TAP2 than the high NCAPG expression group, which may explain the worse prognosis of low NCAPG expression group. Our findings proposed a link of NCAPG expression with tumor stemness.

Although we have integrated and investigated data from different databases, certain limitations remain in our work. First, although bioinformatics provided useful insights into

NCAPG expression in cancer, in vivo and in vitro biological experiments are required to validate our results to facilitate clinical applications. Mechanistic studies are warranted to elucidate the function of NCAPG at cellular and molecular levels. Specifically, the lack of support from prospective clinical studies may be compensated if actual patient data are available. In the future, emphasis will be on the role of NCAPG in STAD progression and regulation of the tumor microenvironment through ex vivo experiments.

In conclusion, we validated that NCAPG is differentially expressed between normal and tumor tissues; the results suggested the correlation between its expression and clinical prognosis in STAD. Our results suggested that the high level of NCAPG expression leads to different prognostic outcomes. With the help of the IRS model and the TIDE algorithm, NCAPG was shown to be effective in predicting immunotherapeutic responses in the STAD cohort. Together, the findings suggest that NCAPG is a valid biomarker for predicting immunotherapeutic responses. These results have implications for elucidating the role of NCAPG in tumorigenesis and progression especially in STAD and provide a reference for achieving precise and personalized immunotherapy in the future.

5. Conclusion

In conclusion, we explored a close relation of NCAPG with tumor microenvironment and prognosis as well as tumor stemness in STAD. We performed a comprehensive evaluation of the pan-cancer potential of NCAPG as a predictive biomarker and identified its significant value in STAD, which may have implications for immunotherapy and is

expected to provide a useful assessment system for clinical application.

Data Availability

The datasets generated and/or analyzed during the current study are available in the GSE66229 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66229), in the GSE84437repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84437), in the GSE26942 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26942), in the GSE13861 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE13861), in the GSE28541 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28541), in the GSE26253 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26253), in the GSE135222 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135222), in the GSE91061 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE91061), in the GSE78220 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78220), and in the GSE84437 repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84437).

Conflicts of Interest

The authors report that there are no competing interests to declare.

Authors' Contributions

All authors contributed to this present work: ZX designed the study, GLC acquired the data, YHW drafted the manuscript, and JKG and JGJ revised the manuscript. All authors read and approved the manuscript.

Acknowledgments

Supplementary Materials

Supplementary 1. FigS1: Kaplan-Meier survival curves of high and low NCAPG expression groups in 33 cancer types.

Supplementary 2. FigS2: Pearson correlation analysis of NCAPG expression with immune score and stromal score.

Supplementary 3. FigS3: the performance of IRS model in GSE84437, GSE26942, GSE13861, GSE28541, and GSE26253 datasets.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] A. Ribas and J. D. Wolchok, "Cancer immunotherapy using checkpoint blockade," *Science*, vol. 359, no. 6382, pp. 1350–1355, 2018.

[3] A. Blum, P. Wang, and J. C. Zenklusen, "SnapShot: TCGA-analyzed tumors," *Cell*, vol. 173, no. 2, p. 530, 2018.

[4] F. Wu, W. Huang, L. Yang, and F. Xu, "MicroRNA-101-3p regulates gastric cancer cell proliferation, invasion and apoptosis by targeting PIM 1 expression," *Cellular and Molecular Biology*, vol. 65, no. 7, pp. 118–122, 2019.

[5] A. Eberlein, A. Takasuga, K. Setoguchi et al., "Dissection of genetic factors modulating fetal growth in cattle indicates a substantial role of the non-SMC condensin I complex, subunit G (NCAPG) gene," *Genetics*, vol. 183, no. 3, pp. 951–964, 2009.

[6] D. Jäger, E. Stockert, E. Jäger et al., "Serological cloning of a melanocyte Rab guanosine $5'$-triphosphate-binding protein and a chromosome condensation protein from a melanoma complementary DNA library," *Cancer Research*, vol. 60, no. 13, pp. 3584–3591, 2000.

[7] Y. Goto, A. Kurozumi, T. Arai et al., "Impact of novel *miR-145-3p* regulatory networks on survival in patients with castration-resistant prostate cancer," *British Journal of Cancer*, vol. 117, no. 3, pp. 409–420, 2017.

[8] M. L. Liang, T. H. Hsieh, K. H. Ng et al., "Downregulation of miR-137 and miR-6500-3p promotes cell proliferation in pediatric high-grade gliomas," *Oncotarget*, vol. 7, no. 15, pp. 19723–19737, 2016.

[9] H. Yan, Z. Li, Q. Shen et al., "Aberrant expression of cell cycle and material metabolism related genes contributes to hepatocellular carcinoma occurrence," *Pathology, Research and Practice*, vol. 213, no. 4, pp. 316–321, 2017.

[10] W. Liu, B. Liang, H. Liu et al., "Overexpression of non-SMC condensin I complex subunit G serves as a promising prognostic marker and therapeutic target for hepatocellular carcinoma," *International Journal of Molecular Medicine*, vol. 40, no. 3, pp. 731–738, 2017.

[11] X. Zhang, M. Zhu, H. Wang et al., "Overexpression of NCAPG inhibits cardia adenocarcinoma apoptosis and promotes epithelial-mesenchymal transition through the Wnt/β-catenin signaling pathway," *Gene*, vol. 766, article 145163, 2021.

[12] B. Song, J. Du, D. F. Song, J. C. Ren, and Y. Feng, "Dysregulation of NCAPG, KNL1, miR-148a-3p, miR-193b-3p, and miR-1179 may contribute to the progression of gastric cancer," *Biological Research*, vol. 51, no. 1, p. 44, 2018.

[13] W. Shen, Z. Song, X. Zhong et al., "Sangerbox: a comprehensive, interaction-friendly clinical bioinformatics analysis platform," *iMeta*, vol. 1, no. 3, 2022.

[14] S. Hänzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.

[15] K. Yoshihara, M. Shahmoradgoli, E. Martínez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, no. 1, p. 2612, 2013.

[16] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.

[17] Y. Liao, J. Wang, E. J. Jaehnig, Z. Shi, and B. Zhang, "WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs," *Nucleic Acids Research*, vol. 47, no. W1, pp. W199–w205, 2019.

[18] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.

[19] P. Blanche, J. F. Dartigues, and H. Jacqmin-Gadda, "Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks," *Statistics in Medicine*, vol. 32, no. 30, pp. 5381–5397, 2013.

[20] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, no. 10, pp. 1550–1558, 2018.

[21] N. Salomonis, P. J. Dexheimer, L. Omberg et al., "Integrated genomic analysis of diverse induced pluripotent stem cells from the progenitor cell biology consortium," *Stem Cell Reports*, vol. 7, no. 1, pp. 110–125, 2016.

[22] B. Ru, C. N. Wong, Y. Tong et al., "TISIDB: an integrated repository portal for tumor-immune system interactions," *Bioinformatics*, vol. 35, no. 20, pp. 4200–4202, 2019.

[23] A. Ponomarev, Z. Gilazieva, V. Solovyeva, C. Allegrucci, and A. Rizvanov, "Intrinsic and extrinsic factors impacting cancer stemness and tumor progression," *Cancers*, vol. 14, no. 4, p. 970, 2022.

[24] H. Suzuki, I. Oda, S. Abe et al., "High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection," *Gastric Cancer*, vol. 19, no. 1, pp. 198–205, 2016.

[25] C. Allemani, H. K. Weir, H. Carreira et al., "Global surveillance of cancer survival 1995-2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2)," *Lancet*, vol. 385, no. 9972, pp. 977–1010, 2015.

[26] J. W. Kim, K. H. Nam, S. H. Ahn et al., "Prognostic implications of immunosuppressive protein expression in tumors as well as immune cell infiltration within the tumor microenvironment in gastric cancer," *Gastric Cancer*, vol. 19, no. 1, pp. 42–52, 2016.

[27] D. Zeng, R. Zhou, Y. Yu et al., "Gene expression profiles for a prognostic immunoscore in gastric cancer," *The British Journal of Surgery*, vol. 105, no. 10, pp. 1338–1348, 2018.

[28] Y. Jiang, Q. Zhang, Y. Hu et al., "ImmunoScore signature," *Annals of Surgery*, vol. 267, no. 3, pp. 504–513, 2018.

[29] P. I. Ribeiro Franco, A. P. Rodrigues, L. B. de Menezes, and M. M. Pacheco, "Tumor microenvironment components: allies of cancer progression," *Pathology, Research and Practice*, vol. 216, no. 1, article 152729, 2020.

[30] T. Ishimoto, H. Sawayama, H. Sugihara, and H. Baba, "Interaction between gastric cancer stem cells and the tumor microenvironment," *Journal of Gastroenterology*, vol. 49, no. 7, pp. 1111–1120, 2014.

[31] D. B. Johnson, R. J. Sullivan, and A. M. Menzies, "Immune checkpoint inhibitors in challenging populations," *Cancer*, vol. 123, no. 11, pp. 1904–1911, 2017.

[32] M. B. Atkins, J. I. Clark, and D. I. Quinn, "Immune checkpoint inhibitors in advanced renal cell carcinoma: experience to date and future directions," *Annals of Oncology*, vol. 28, no. 7, pp. 1484–1494, 2017.

[33] J. M. Lee, M. H. Lee, E. Garon et al., "Phase I trial of intratumoral injection ofCCL21gene–modified dendritic cells in lung cancer elicits tumor-specific immune responses and CD8+ T-cell infiltration," *Clinical Cancer Research*, vol. 23, no. 16, pp. 4556–4568, 2017.

[34] P. S. Hegde, V. Karanikas, and S. Evers, "The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition," *Clinical Cancer Research*, vol. 22, no. 8, pp. 1865–1874, 2016.

[35] S. K. Bunt, L. Yang, P. Sinha, V. K. Clements, J. Leips, and S. Ostrand-Rosenberg, "Reduced inflammation in the tumor microenvironment delays the accumulation of myeloid-derived suppressor cells and limits tumor progression," *Cancer Research*, vol. 67, no. 20, pp. 10019–10026, 2007.

[36] T. M. Malta, A. Sokolov, A. J. Gentles et al., "Machine learning identifies stemness features associated with oncogenic dedifferentiation," *Cell*, vol. 173, no. 2, pp. 338–354.e15, 2018.

[37] Z. Ye, M. Zheng, Y. Zeng et al., "Bioinformatics analysis reveals an association between cancer cell stemness, gene mutations, and the immune microenvironment in stomach adenocarcinoma," *Frontiers in Genetics*, vol. 11, article 595477, 2020.

[38] J. Agudo, E. S. Park, S. A. Rose et al., "Quiescent tissue stem cells evade immune surveillance," *Immunity*, vol. 48, no. 2, pp. 271–85.e5, 2018.