Research Paper

ABSTRACT

INTRODUCTION

Prostate cancer (PCa), a malignant cancer, still one of the top ten cancers in the world, the most commonly diagnosed in the USA and ranked sixth in morbidity and seventh in mortality in China [1, 2]. At present, therapies surgery, androgen-deprivation therapy (ADT), radiation (RT), ablative , chemotherapy, and immune check-point [3]. many early clinical PCa patients with low-risk or intermediate-risk could be greatly cured by performing surgery, ablative therapies, or RT, they might possibly develop biochemical recurrence of PCa [4, 5]. In addition, more advanced PCa patients with high risk who received ADT and chemotherapy finally progressed to refractory castration-resistant prostate cancer (CRPC), which was considered to be closely associated with the poor clinical prognostic factor androgen receptor (AR) in PCa [3, 6]. Immune checkpoint) therapy, as a newly treatment, to be an efficient method to relieve CRPC [7]. Unfortunately, it was reported that some immune therapies targeting T-cell immune including CTLA-4, PD-L1 and PD-1 been demonstrated to be significantly efficient in PCa, which might be connected with the poor tumor immune

microenvironment in PCa [8]. According to literature review, PCa considered a cold tumor with low levels of infiltrating T cells, defective function of antigen presenting cells ( ) and high levels of infiltrating immunosuppressive cells such as MDSCs and Tregs, the latter secrete cytokines to suppress immune function and inhibit immune activation in tumor microenvironment (TME), contributing to the immune escape of PCa cells [9]. Therefore, new methods are needed to accurately predict the prognosis and immune escape of PCa.

Integrins, known as a kind of transmembrane - composed of two different heterodimers 18 and 8 integrin polypeptides, which with various ligands to produce many pathophysiological effects including adhesion, [10]. Among others, integrins and were to be overexpressed in PCa patients and were connected with tumor growth, angiogenesis and metastasis in PCa [11]. Rubenstein, C. S. et al. demonstrated that PCa cells

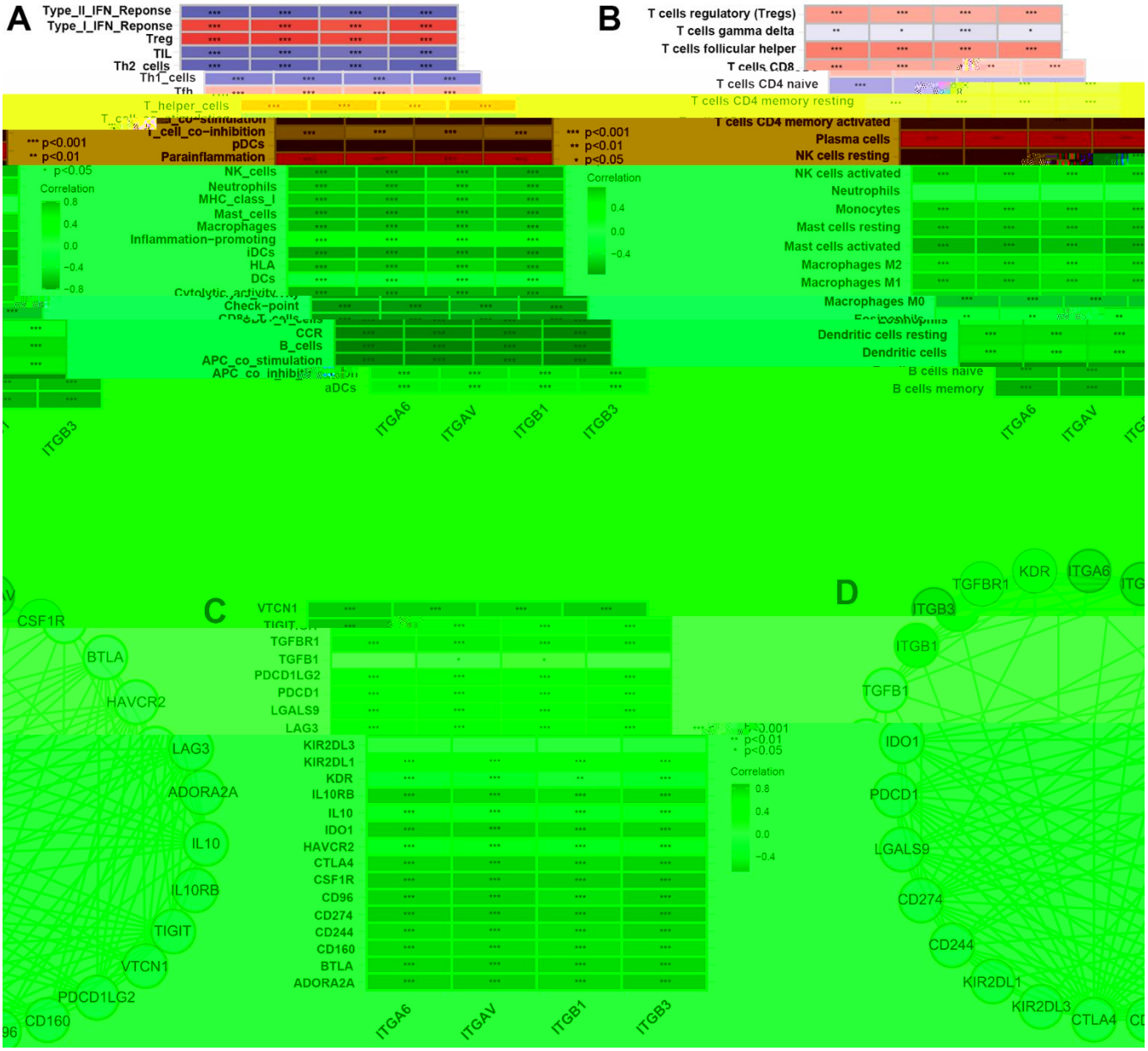
invasive network on laminin-containing and invaded smooth muscle both *in vitro* and *in vivo* and integrin enhanced the intercellular biophysical properties in PCa [12].
regarded as a potential biomarker for early
neuroendocrine PCa, and according to literature, integrin 1 strongly connected with PCa
after radical surgery, suggesting integrins a crucial role in the invasion, progression and prognosis of PCa [13, 14]. Furthermore, integrins and integrin ligands also vital roles in infiltration and activation of T cells and tumor immune microenvironment, which a dual effect of promotion and in cancers [15]. reported the potential association abnormally expressed integrins PCa nevertheless, comprehensive analysis the value of integrin / prognosis and immune escape in PCa still inadequate [16].

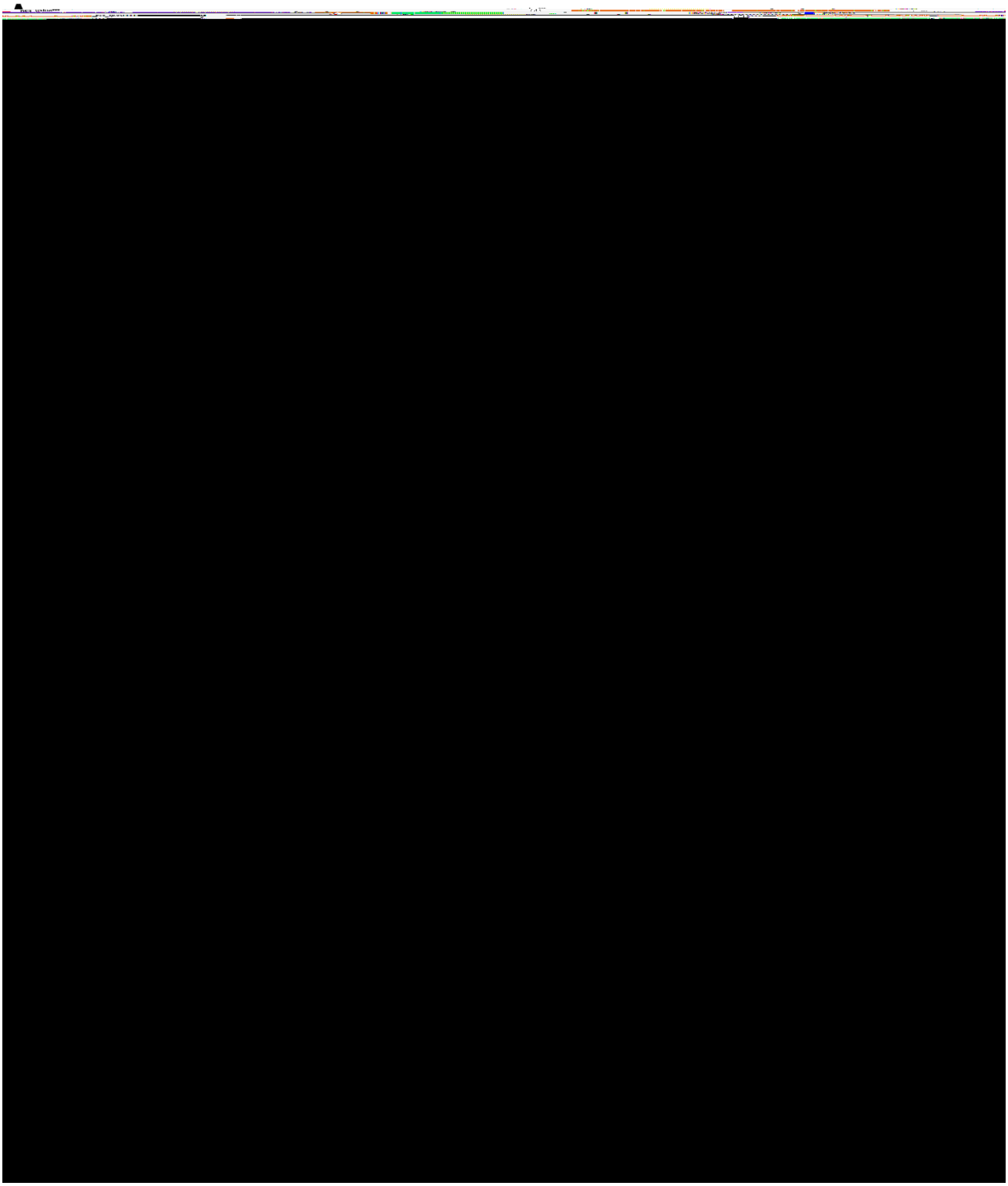
In this work, we intend to construct a risk model based on integrin / to effectively predict prognosis and immune escape in PCa, and these findings may also provide new potential targets for precision treatment of PCa.

Integrin / were closely associated with poor immunity in PCa patients

, we systematically analyzed the correlations among integrin with immune infiltration, immune functions and immunosuppressive genes. Interestingly, we found that integrin negatively associated with immune-4(i)5

high expression levels of ITGB1 was significantly lower than the samples with low expression levels of ITGB1 (Figure 2C). Likewise, the RFS for the samples with high expression levels of ITGAV and ITGB1 markedly lower than the samples with low expression levels of ITGAV and ITGB1, which might indicate that integrin could be a potential prognostic element in PCa (Figure 2D, 2E). Furthermore, PCa patients who high levels of integrin ITGAV, ITGA6, ITGB1 and ITGB3 higher proportion of biochemical recurrence patients with low expression levels of integrin (Figure 2F 2I). , we further investigated the correlations among the expression levels of ITGAV, ITGA6, ITGB1 and ITGB3 the expression levels of AR in PCa patients. According to literature reports, androgen receptor (AR) one of the prognostic factors in PCa, to the progression and recurrence of PCa. Surprisingly, there were strong correlations among the ITGAV, ITGA6, ITGB1 and ITGB3 with AR in PCa patients, whose levels for cor-





0.759, 0.650, 0.575, 0.448, respectively (Figure 2J 2M). In summary, these results suggested that integrin

PCa.

A prognostic risk model based on integrin / for PCa patients

, we divided the PCa patients into two (C1 and C2) according to the average expression levels of whole integrin (ITGAV, ITGA6, ITGB1, ITGB3) in TCGA cohort and we performed gene cluster analysis for two to identify significant DEGs based on the whole genetic transcriptomes of 496 PCa patients in TCGA dataset, and 13994 DEGs were identified in the two clusters (Figure 3A).
clinical characteristics of these PCa patients was also manifested the heatmap. Next, 51 DEGs identified as prominently affecting biochemical recurrence of PCa by conducting univariate analysis, including 33 DEGs (hazard ratio >1) increased the risk of biochemical recurrence and 18 DEGs (hazard ratio <1) decreased this risk in the TCGA cohort (Figure 3B). Subsequently, we intersected these 51 DEGs with the DEGs acquired from the GEO dataset and successfully identified eight biochemical recurrence-related genes (ASF1B, INSM2, POU4F2, MT1B, NCR1, KRTAP10-5, PCDHA13 and KIR3DL1) in both TCGA and GEO cohorts. Next, we constructed a risk model using the value based on the eight genes by applying LASSO regression analysis (Figure 3C, 3D). In addition, we calculated the risk score of these 8 biochemical recurrence-related genes by running the formula: Risk score= (0.0741808537008443 * expression of ASF1B) + (0.0584797685631388 * expression of INSM2) + (-0.0665901093358763 * expression of POU4F2) + (0.152621619239078 * expression of MT1B) + (0.00167204457436465 * expression of NCR1) + (-0.0181031499555924 * expression of KRTAP10-5) + (0.151620509953615 * expression of PCDHA13) + (-0.0816400681584183 * expression of KIR3DL1).

Based on the median risk score, PCa patients in the TCGA cohort and GEO cohort, which had complete biochemical recurrence information, were objectively divided into the HR group and LR group (Supplementary Figure 1A). The genes ASF1B, INSM2, MT1B, NCR1 and PCDHA13 were highly expressed in the HR group, while the genes POU4F2, KRTAP10-5 and KIR3DL1 were highly expressed in the LR group in both the TCGA and GEO cohorts (Figure 3E, 3F). To further confirm the relationship between these genes and biochemical recurrence, we selected 5 cases of prostate cancer and 5 cases of prostate cancer for RNA extraction. results showed that the expression levels of ASF1B, INSM2, MT1B, NCR1 and PCDHA13 in biochemically recurrent PCa were higher than those in non-biochemically recurrent PCa, while POU4F2, KRTAP10-5 and KIR3DL1 were significantly lower in biochemically recurrent PCa than in non-biochemically recurrent PCa (Figure 3G). Additionally, we were able to find that there were more dead samples that were regarded as the HR group located in the area of higher risk scores with increasing risk scores of patients. In contrast, the other survivors who were regarded as the LR group were more likely to be located in areas with lower risk scores, and they had longer survival times than the HR group (Supplementary Figure 1B). Otherwise, we applied principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) analysis to demonstrate that samples in the HR and LR groups could be greatly separated (Supplementary Figure 1C, 1D).

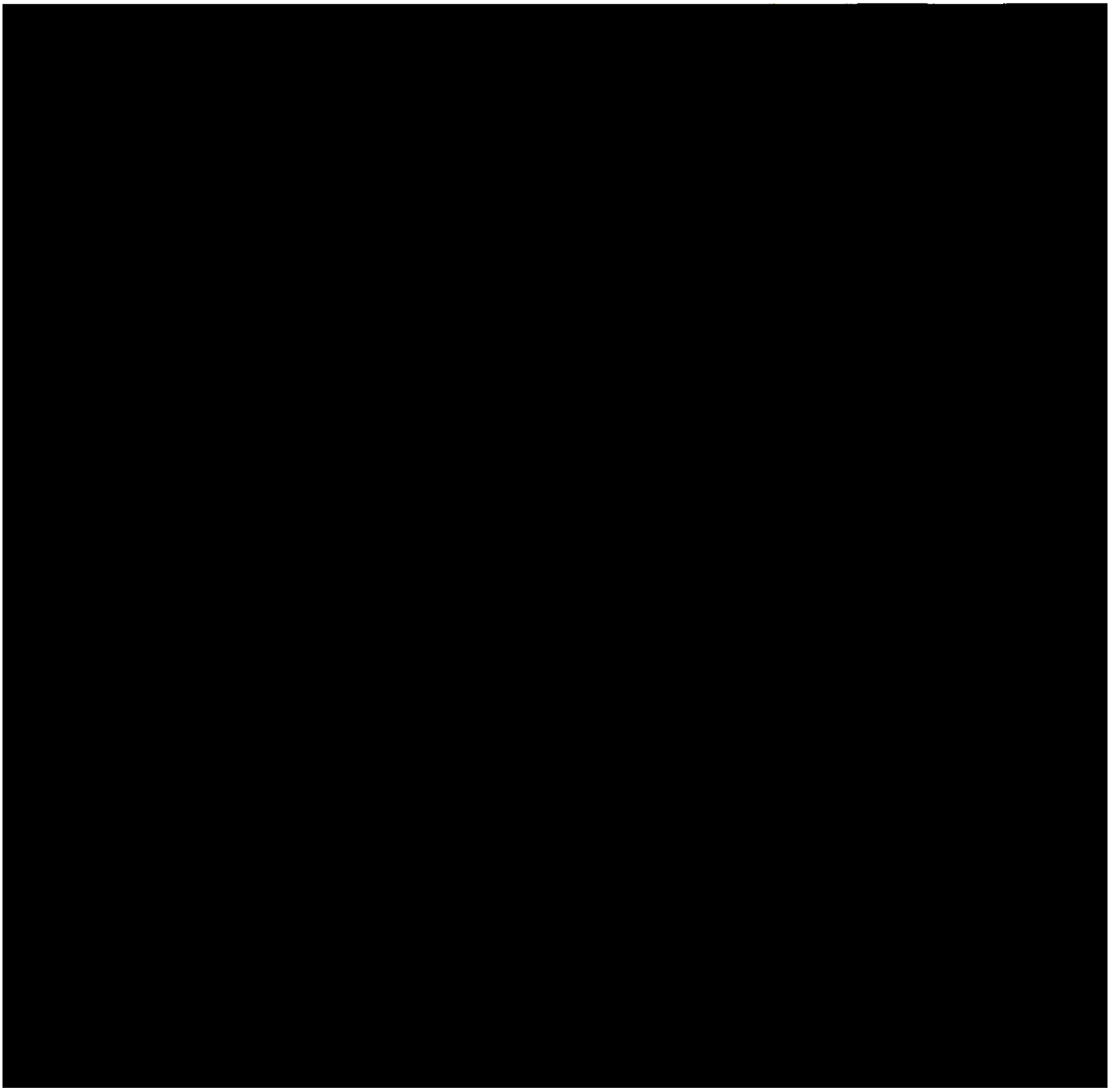
Furthermore, we carried out Kaplan Meier analysis for biochemical recurrence between the two groups, and the results revealed worse biochemical recurrence probabilities for PCa patients in the HR group than in the LR group (Figure 4A). ROC analysis was performed to prove the predictive ability of our model for biochemical recurrence, and the area under the ROC curve (AUC) was 0.716, 0.764, and 0.832 for 1, 3, and 5 years, respectively (Figure 4D). Likewise, worse progression-free survival (PFS) and OS were observed in the HR group than in the LR group (Figure 4B, 4C). The AUC values for PFS were 0.950 at 1 year, 0.906 at 3 years, and 0.869 at 5 years, while the AUC values for OS were 0.992 at 1 year, 0.906 at 3 years, and 0.912 at 5 years (Figure 4E, 4F). Moreover, univariate and multivariate showed that risk score was an independent risk factor for PCa patients (Figure 4G, 4H). These evidences illustrate that our risk model was powerful the prognosis PCa patients.

The HR risk group predicted poor clinical features of PCa patients

On the other hand, we the correlations of the two risk groups with clinical features in PCa patients. results that the percent weight of clinical and pathologic T3-stage and T4-stage in the HR group was higher than the LR group (40% vs 15% and 95% vs 80%), and the risk score of clinical and pathologic T3-stage and T4-stage was higher than clinical and pathologic T1-stage and T2-stage,

respectively (Figure 5A 5D), suggesting more patients suffered advanced PCa in HR group. , the percent weight of fustat-1 (the dead patients) in HR group was markedly higher than LR group (74% vs 41%), the same result risk score in two groups (Figure 5E, 5F). In addition, the percent weight of PSA (value >4) which could greatly identify PCa in patients was dramatically higher in HR group than LR group (15% vs 2%), and the risk score of PSA (value >4) was higher than PSA (value <4) in two groups as well (Figure 5G, 5H).

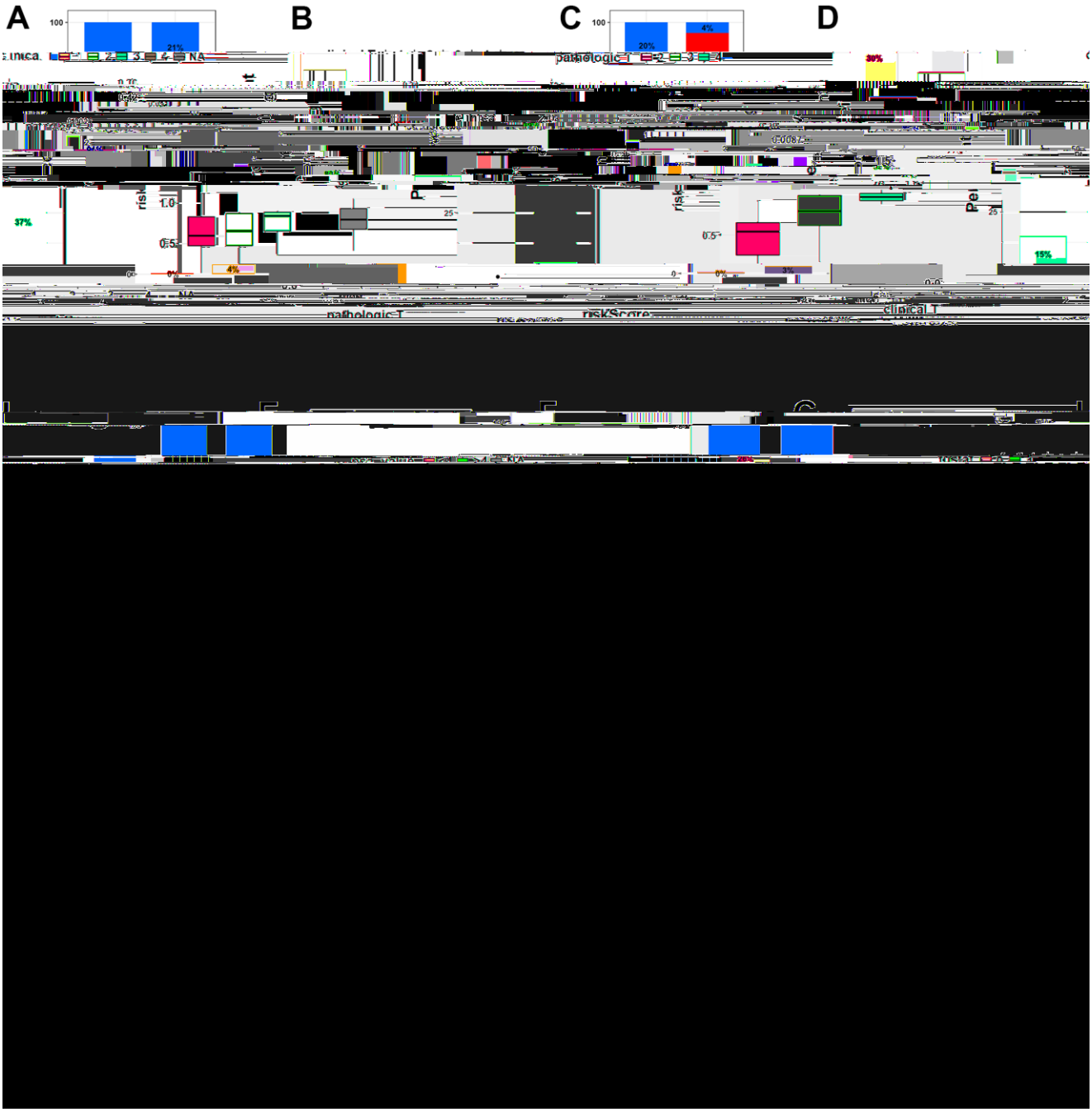
Moreover, survival analysis was performed for the two risk groups under subgroups of clinical of PCa patients. The results showed that PCa patients with clinical T1-stage, T2-stage



and T3-stage probably more mortality in HR group than LR group (Figure 5I 5K). Similarly, the patients with Gleason score 7-9, PSA (value >4 and <4), pathologic stages of T2-T3 and N0-N1 both lower survival probability in the HR group than in the LR group (Supplement Figure 3A 3I). strongly demonstrated the HR group could well predict poor survival in PCa patients.

Genetic mutation status and functional enrichment analysis for the two risk groups

In the TCGA cohort, the GO enrichment analysis showed that GOBP: NIK (NF- - signaling, GOBP: ras protein signal transduction, GOBP: canonical signaling pathway and GOBP: positive regulation of interleukin-10 production was



markedly enriched in the HR group LR group. , according to GEO cohort, GOBP: transforming growth factor beta1 production and GOMF: heat shock protein binding was more enriched in HR group than LR group. - - and FC receptor were more significantly enriched in HR group than LR group both TCGA and GEO cohorts (Figure 6A, 6B and Supplementary Tables 1, 2). In addition, KEGG enrichment analysis showed that KEGG: T-cell and B-cell receptor signaling pathway, KEGG: P53 signaling pathway, KEGG: wnt signaling pathway, TGF beta signaling pathway, chemokine signaling pathway, KEGG: cell cycle and KEGG: MAPK signaling pathway highly enriched in HR group LR group in TCGA or GEO cohorts (Figure 6C, 6D and Supplementary Tables 3, 4). Based on these results, we can note that most of the functions or pathways enriched in HR group are associated with poor outcomes. Therefore, these evidences might reveal the potential mechanisms for poor prognosis in the HR group.

To evaluate the difference genetic mutation status in two risk groups, we Maftools to identify the whole gene mutation in 91 PCa patients complete biochemical recurrence data from TCGA dataset, of which 20 most frequently mutated genes shown in Figure 7A, 7B. We noticed higher frequently mutated ratio in HR group LR group (75.56% vs 58.70%). Moreover, the TP53 and KMT2D were more often mutated in HR group than in LR group (18% vs 15% and 9% vs 7%). However, the gene SPOP rates were higher in LR group than in HR group (20% vs 9%), which could be attributed to genetic hetero-geneity. Collectively, these mutated genes might partially to the poor clinical prognosis of PCa in HR group.

The HR group predicted a higher risk of immune escape for PCa patients

Moreover, we further investigated the of immune cells infiltrated the TME of PCa between two groups in TCGA samples (Figure 7C). results showed lower of , B cells, $CD8^+$ T cells, macrophages, NK cells, Th cells and TILs in HR group than LR group. negatively regulated immune cells such as Treg cells were higher in the HR group than LR group. To further investigate the influences of immune cells our patients was performed for PFS, OS and biochemical recurrence in TCGA cohort. As expected, the high expression levels of $CD8^+$ T cells, NK cells and exhibited better PFS, OS or biochemical recurrence, while the high expression levels of exhibited poor PFS and OS (Supplementary Figure 2A 2H). These results were consistent with conventional wisdom.

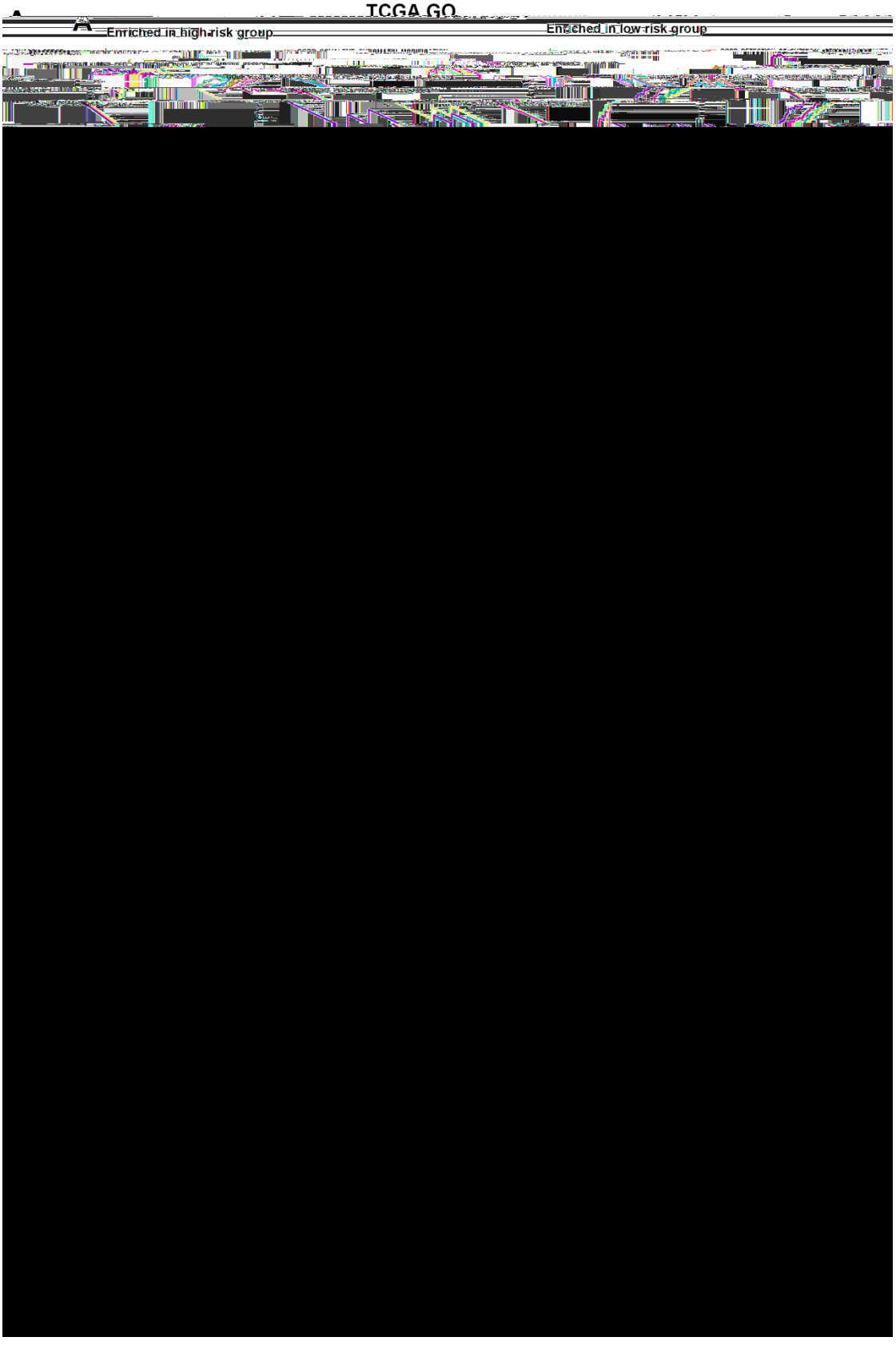
Additionally, immune functions were also investigated in the HR and LR groups, and the results showed that poor immune functions including aDCs, B cells, cells, NK cells and macrophages, and better Treg functions were found in the HR group than LR group. functions such as APC - , T-cell - were also lower in the HR group than in the LR group (Figure 7D). These evidences illustrate worse immune functions in the HR group which might contribute to tumor immune escape in PCa. Next, we compared the expression of immune checkpoints in risk groups and found that the HR group had higher expression of PD-1 (CD274), PD-L1 (PDCD1), and CTLA-4 than the LR group, which might lead to higher risk of immune escape in PCa (Figure 7F 7H). In addition, we scored the risk of immune exclusion in our patients surprisingly, we found that the immune exclusion score was higher in the HR group than in the LR group (Figure 7E). These findings suggested the HR risk group might suffer a higher risk of immune escape in PCa patients.

Integrins are ubiquitous heterodimeric transmembrane glycoprotein receptors that interact with ligands in cells and the extracellular matrix and mainly act as bidirectional signaling proteins to regulate various physiological functions of cells, such as mediating cell adhesion and promoting tumorigenesis, progression, and metastasis in tumors [17, 18]. The biological

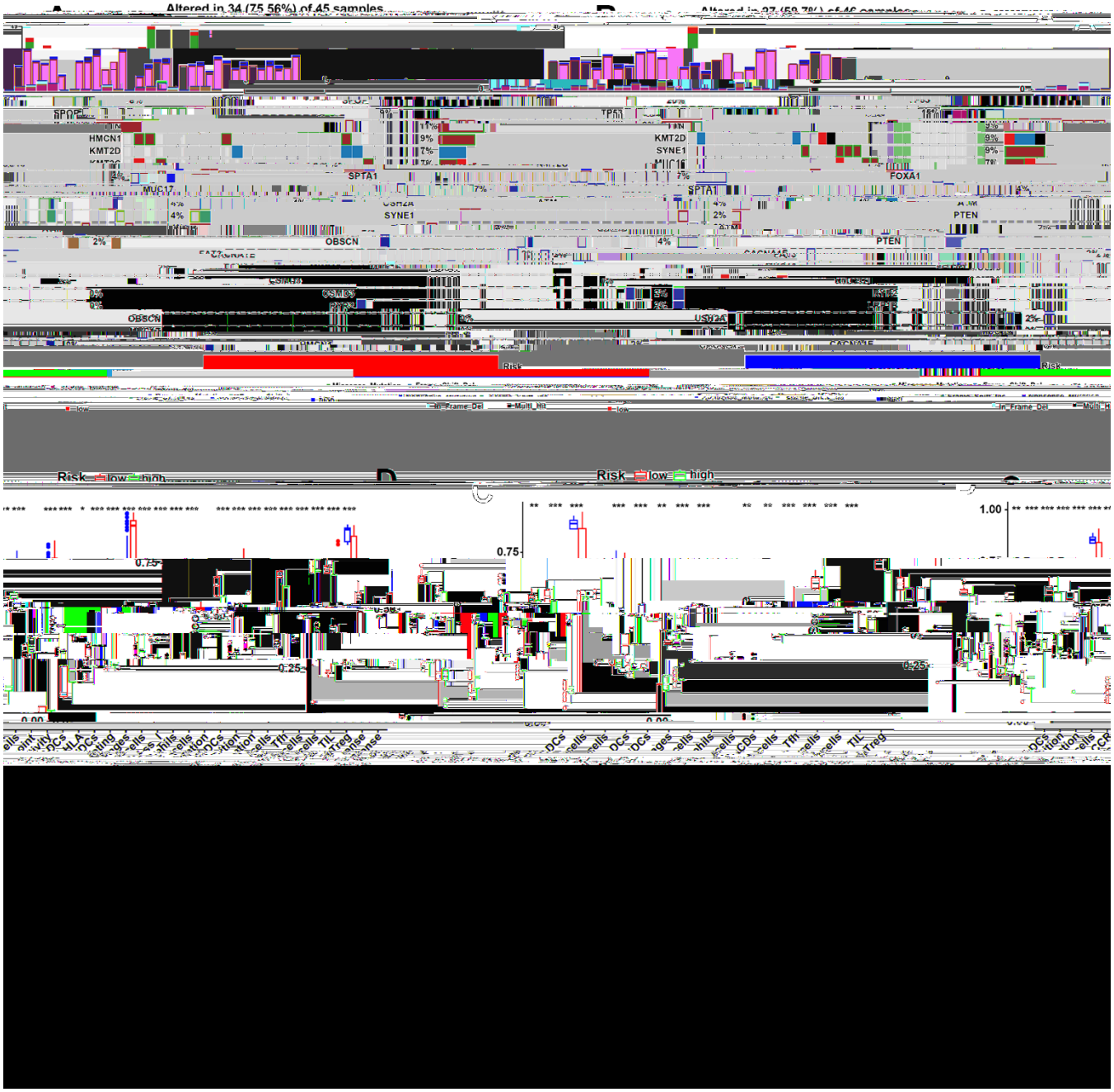
particularly important and have been extensively

adhesion to the extracellular matrix (ECM) by targeting RGD (Arg-Gly-Asp) in fibronectin and regulates MMP (especially MMP-2 and MMP-9) expression through the PI3K signaling pathway to hydrolyze collagen in the

participates in almost the entire growth processes of PCa, especially contributing to progression, angio-genesis and metastasis by combining with ligands expressing the RGD sequence, for instance, bone sialoprog0 G[(t)-5(h)] TJETQq0.00000912 0 6throughgh ttastasis


A
TCGA GO
Enriched in high-risk group
Enriched in low-risk group

kine C-C motif ligand 2 (CCL2) was proven to increase migration of PCa [21]. In addition, a recent study invade surrounding nerves and supported PCa bone metastasis when laminin was combined with integrin which was stimulated by active androgen receptor (AR) in castration-resistant prostate cancer (CRPC), could promote the accumulation of Bnip3 to prevent the apoptosis of CRPC cells [23]. Moreover, our previous study revealed that endostatin 33 peptide is



by inhibiting the PI3K-Akt signaling pathway in prostate cancer [24]. In this work, we integrated the genetic characteristics of integrins
focused more on the derivative value of integrin

critical value in predicting OS, PFS, and biochemical recurrence in PCa patients. This evidence indicates that

indicator of the diagnosis and prognosis of PCa and that the genes that construct the model might be new therapeutic targets.

On the other hand, several studies have suggested a close relationship between integrins and tumor immunity. Bagati A et al. confirmed that breast cancer could escape $CD8^+$ T-cell attack through the integrin 6-TGF -SOX4 pathway, which significantly decreases tumor sensitivity to $CD8^+$ T cells [25]. According to a literature review, integrin especially upregulates TGF beta, contributing to promoting angiogenesis and immune suppression in cancers [26]. Integrin was also proven to promote immune escape by regulating the interferon signaling pathway and increasing PD-L1 expression in cancers [19, 27]. Additionally, Yang H et al. showed that ITGB1 suppressed T-cell function, contributing to immune evasion and a low response to immune checkpoint treatment in pancreatic cancer [28]. Despite this, the effect of integrins on immune function in PCa is not well understood. In our work, we found that integrin were negatively correlated with most immune cell infiltration and immune functions in PCa patients, suggesting that integrin might be a reason for immune evasion of PCa, which is worth further discussion in follow-up studies.

In our study, we identified eight biochemical recurrence-related genes (ASF1B, INSM2, POU4F2, MT1B, NCR1, KRTAP10-5, PCDHA13, and KIR3DL1) to establish a risk model. The expression of ASF1B, INSM2, MT1B, and PCDHA13 was markedly upregulated in the HR group. ASF1B (antisilencing function 1B) is a histone H3-H4 chaperone protein that regulates the functions of chromatin in cells [29]. Han G et al. demonstrated that ASF1B was overexpressed in PCa tissues, which was closely connected with the low OS of PCa patients, since ASF1B could promote tumorigenesis through the ASF1B-PI3K\AKT signaling pathway mediating the cell cycle in PCa [30]. Similarly, it was reported that ASF1B was significantly correlated with poor prognosis in several other cancer patients [31, 32]. Cao H et al. demonstrated that the upregulated expression of INSM2 promoted tumorigenesis and progression by regulating lipid metabolism in neuroblastoma and led to poor prognosis in neuroblastoma patients [33]. MT1B, known as an important isoform of metallothioneins (MTs), was reported to participate in the regulation of copper-zinc homeostasis, and MTs play an important role in tumorigenesis, angiogenesis, metastasis, proliferation and immunomodulation of cancers [34]. Wang KH et al. illustrated that methylated PCDHA13 and PCDHA4 were closely associated with the severity and progression of cervical carcinoma, and PCDHA13 and PCDHA3 were considered screening biomarkers with more specificity and equal sensitivity when combined with HPV to test cervical cancers compared with the HPV test alone [35]. Together, these evidences show that the overexpressed genes in the HR group were strongly associated with the poor prognosis of cancers, which supported our results that the HR group could predict the prognosis of PCa in our model.

As recently described, immune checkpoint inhibitors (ICIs), including CTLA4 inhibitors and PD1 and PD-L1 inhibitors, have been approved for treating many cancers, such as kidney cancer, melanoma, urothelial cancer and lung cancer, and have achieved good outcomes [36]. However, according to the literature reviews, the monotherapy and combination of ICIs showed limited benefits with low levels of therapeutic responses in CRPC, without a prominent survival improvement, while the combination therapies of ICIs with vaccines, androgen receptor targeting inhibitors (ARTi), chemotherapeutic drugs, poly ADP ribose polymerase (PARP) inhibitors or tyrosine kinase inhibitors (TKI) exhibited a prospective method for the treatment of CRPC [37 39]. Interestingly, Dong M et al. proved that the complex consisting of CUL3 and SPOP proteins could lead to the degradation of PD-L1 through ubiquitination, inhibiting the immune evasion of ovarian cancers [40]. This evidence strongly indicates the importance of seeking new immunotherapy targets in cancer treatment. More importantly, regulating the integrin pathway and expression might enhance the antitumor effects of ICIs in cancers, and newly developed antibodies aimed at integrins are worth evaluating in PCa in the future [41, 42]. In our study, we developed an integrin -based risk model that can accurately predict immune cell infiltration, immune cell function, immune exclusion score and immune checkpoint levels (PD-1, PD-L1 and CTLA-4) in PCa patients. These results indicate that integrin our model are expected to be novel molecular targets for immune escape therapy of PCa.

In conclusion, we established a prognostic risk model based on integrin / which revealed the

TP53 mutation and immune escape of PCa. These evidences might provide medical therapeutic targets for PCa patients.

Data sources

The data of genetic transcriptomes, tumor mutational burden (TMB), SNPs and clinical characteristics for 495 PCa samples were acquired from the TCGA dataset (https://tcga-data.nci.nih.gov/tcga/), and the genetic transcriptomes and clinical features for 220 PCa cases were downloaded from the GEO dataset (http://www.ncbi.nlm.nih.gov/geo/). The results of immunohistochemistry (IHC) of prostate tissues and PCa were acquired from the Human Protein Atlas (HPA) dataset (https://www.proteinatlas.org/). The expression level of integrin v 3/ 6 1 in each clinical PCa patient was determined by the mean of the expression levels of the integrin subunits ITGAV, ITGA6, ITGB1, and ITGB3.

Immune interaction network

We utilized CIBERSORT to information immune infiltrating cells and immune functions PCa patients. Pearson analysis was performed for integrin cells, immune functions and immunosuppressive genes PCa patients. interaction network (PPI) for integrin with immunosuppressive genes was formed by applying STRING (https://string-db.org/) and Cytoscape software.

Identification of differentially expressed genes

, the expression data in both TCGA and GEO cohorts were standardized with the formula log2(x+1) and we intersected TCGA genetic matrix with GEO genetic matrix to the expression data of same genes in both TCGA and GEO cohorts. Then, the data of differentially expressed genes (DEGs) were finally identified by applying edge R package with FDR < 0.05 and 0.585 in both TCGA and GEO cohorts.

Establishment of the prognostic risk model

We intersected biochemical recurrence-related genes in both TCGA and GEO and eight biochemical recurrence-related genes were finally identified, which used to establish our prognostic risk model by applying LASSO regression analysis. In our work, the risk score was calculated by $\sum_{i}^{8} Ai \times Bi$ (A: coefficients, B: gene expression level), which was utilized to separate our samples into two groups (HR and LR group) in both the TCGA training cohort and GEO testing cohort.

Functional analysis

In this work, we first performed gene differential analysis of the HR group and LR group and identified 10407 and 1922 significantly different genes in the TCGA dataset and GEO dataset, respectively. These genes were used in the Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) to run gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses for the two risk groups. For gene set enrichment analysis (GSEA), TCGA and GEO transcriptome data and risk files were input into R software. The program performs enrichment analysis by sorting the expression profile data, calculating enrichment scores, estimating significance levels, and performing multiple hypothesis testing.

Statistical analysis

We employed one-way ANOVA and t tests to compare two risk groups, and comparisons of two or more component percent rates were made using the chi-square test. In addition, heatmaps, forest graphs, box plots and receiver operating characteristic (ROC) curves were completed by R software (version 3.5.1). Waterfall curves were completed by MAF tools and R software. Otherwise, Kaplan Meier analysis was performed to estimate the survival and biochemical recurrence of PCa patients in both the TCGA cluster and two risk groups in the model. We processed all statistical analyses by utilizing SPSS 19.0 software (SPSS. Inc., Chicago, IL, USA) or R software. In our study, a P-value < 0.05 was statistically significant.

Data availability statement

All data in this study are available in the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) and GEO data portal (http://www.ncbi.nlm.nih.gov/geo/).

PCa: prostate cancer; SNPs: single nucleotide polymorphisms; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; HPA: Human Protein Atlas; LASSO: least absolute shrinkage and selection operator; HR group: high-risk group; LR group: low-risk group; AUC: Area under curve; ADT: androgen deprivation therapy; RT: radiation therapy; CRPC:

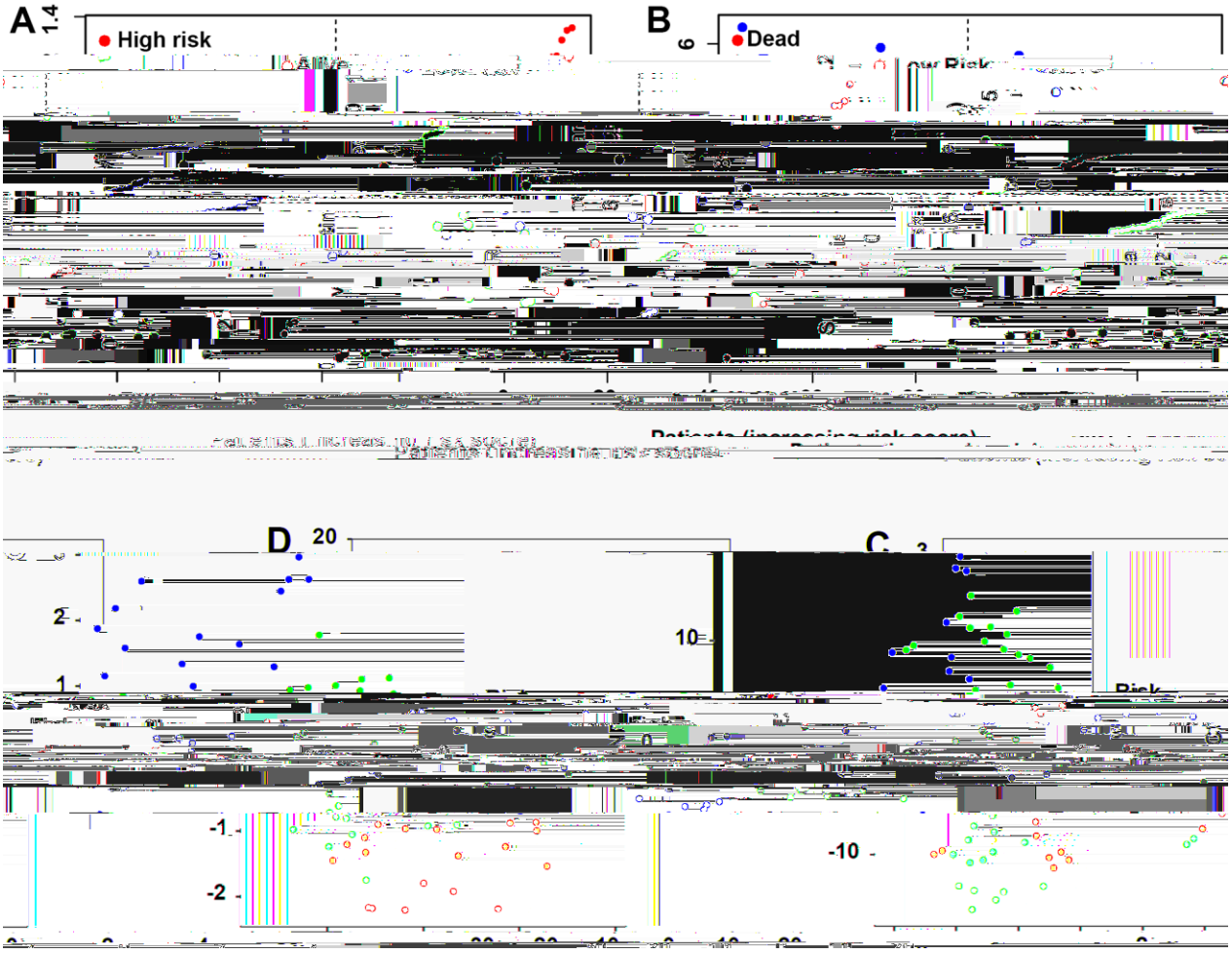
castration-resistant prostate cancer; AR: Androgen receptor; ICIs: immune checkpoint inhibitors; APC: antigen presenting cells; TME: tumor micro-environment; TMB: tumor mutational burden; PPI: DEGs: differentially expressed genes; GSEA: Gene set enrichment analysis; DAVID: Database for Annotation, Visualization and Integrated Discovery; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ROC: Receiver operating characteristic; TILs: Tumor-infiltrating lymphocytes; PSA: Prostate-specific antigen; OS: Overall survival; RFS: Relapse-free survival; PFS: Progression-free survival; PCA: principal component analysis; t-SNE: t-distributed stochastic neighbor embedding; MTs: metallothioneins; recurrence-related genes.

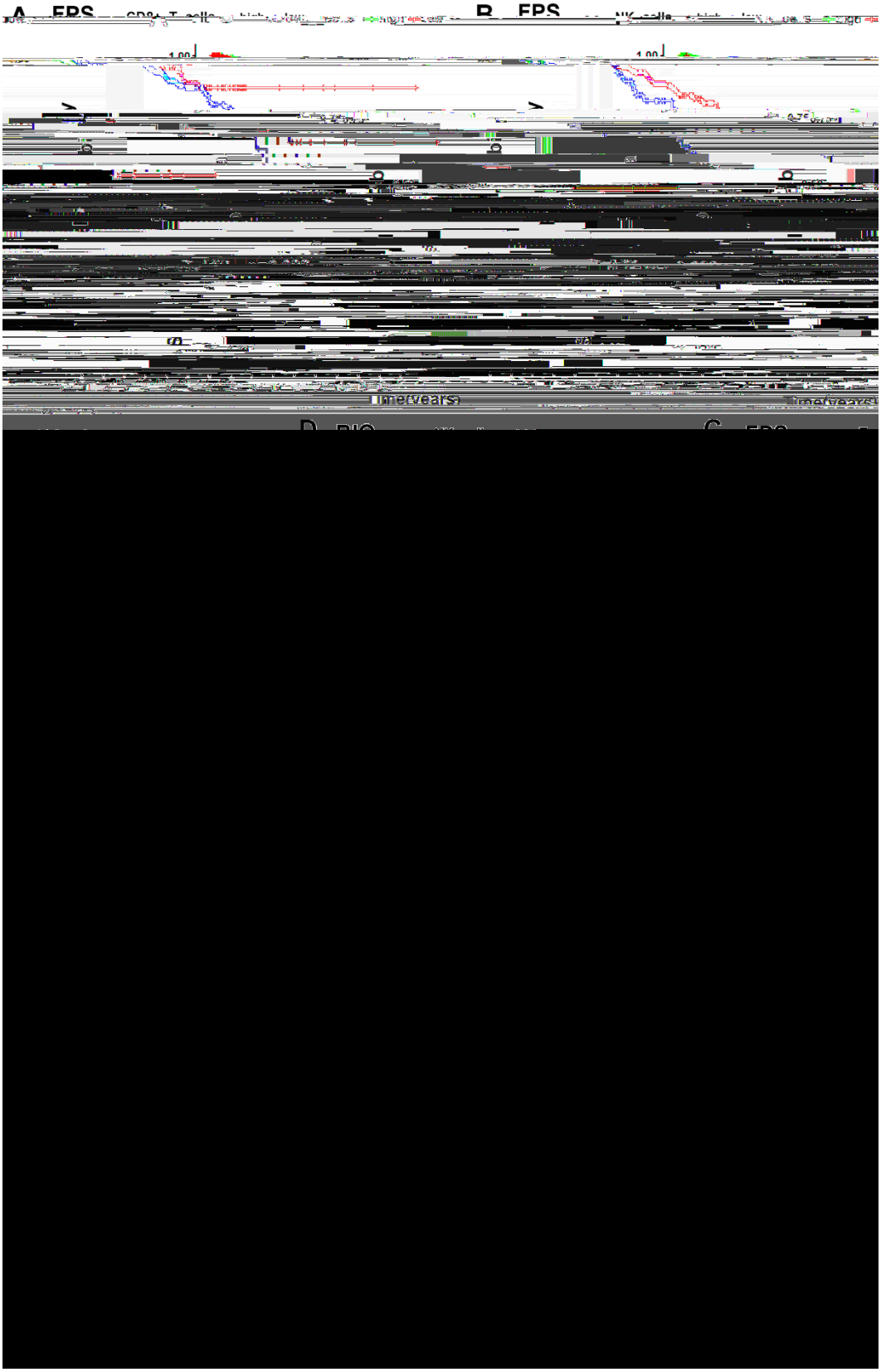
R.M.H. and D.B.C. designed this study and completed the final manuscript. L.Y. analyzed the data, and J.B. interpreted these data. L.Y., J.B. and H.J.X. wrote the manuscript together, and Y.L. and H.J.X. revised the original manuscript. W.J.S., Z.L., P.Z.Q. and W.J.S. participated in all data gathering, visualization, and detection. All authors reviewed and approved this final manuscript.

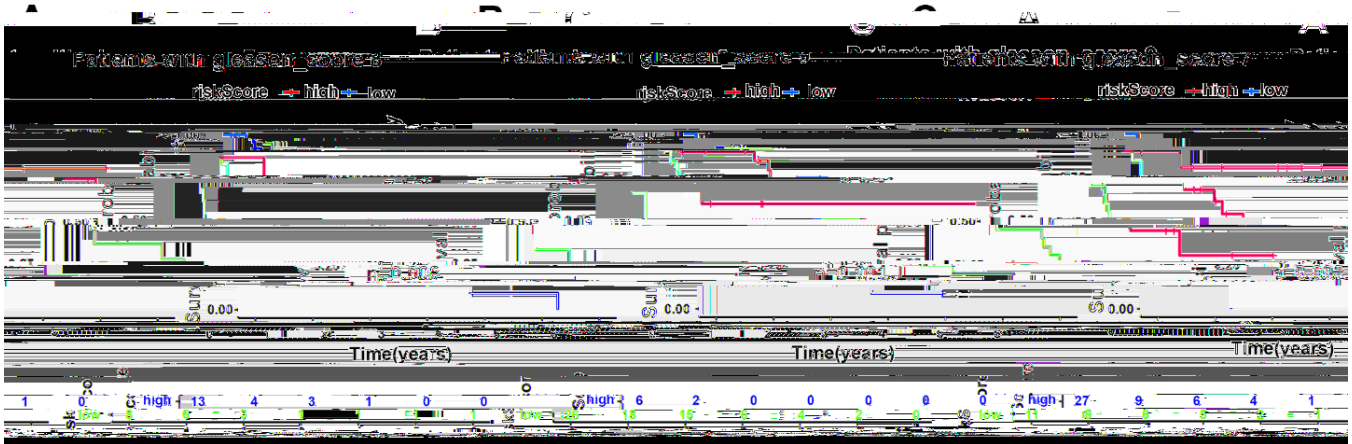
We thank the TCGA and GEO for their data .

The authors declare that they have no conflicts of interest.

This study was supported by the Horizontal Project of Heilongjiang Renxin Medical Assistance Foundation, Heilongjiang Charity Fund General Association, Natural Science Foundation of Heilongjiang Province (No. LH2019H030 to Ming-hua Ren) and National Natural Science Foundation of China (NSFC 82002680 to Bei-chen Ding).


A
High risk
B
Dead
D
20
2
1
-1
-2
C
10
-10




riskScore high low
0.00
Time(years)
high
low

Please browse Full Text version to see the data of Supplementary Tables 1 4.