CELL CYCLE
2023, VOL. 22, NOS. 21–22, 2392–2408
https://doi.org/10.1080/15384101.2023.2293594

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER



# Identification of novel potential homologous repair deficiency-associated genes in pancreatic adenocarcinoma via WGCNA coexpression network analysis and machine learning

Chun Liu[a*], Jingyun Fang[b*], Weibiao Kang[c], Yang Yang[c], Changjun Yu[c], Hao Chen[d], Yongwei Zhang[e], and Huan Ouyang[c]

[a]Department of General surgery, The People's Hospital of Chizhou, Chizhou, Anhui Province, China; [b]Department of Nursing, The People's Hospital of Chizhou, Chizhou, Anhui Province, China; [c]Department of General surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui Province, China; [d]Department of Emergency Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui Province, China; [e]Department of general surgery, Anqing First People's Hospital, Anqing, Anhui Province, China

**ABSTRACT**

Homologous repair deficiency (HRD) impedes double-strand break repair, which is a common driver of carcinogenesis. Positive HRD status can be used as theranostic markers of response to platinum- and PARP inhibitor-based chemotherapies. Here, we aimed to fully investigate the therapeutic and prognostic potential of HRD in pancreatic adenocarcinoma (PAAD) and identify effective biomarkers related to HRD using comprehensive bioinformatics analysis. The HRD score was defined as the unweighted sum of the LOH, TAI, and LST scores, and it was obtained based on the previous literature. To characterize PAAD immune infiltration subtypes, the "ConsensusClusterPlus" package in R was used to conduct unsupervised clustering. A WGCNA was conducted to elucidate the gene coexpression modules and hub genes in the HRD-related gene module of PAAD. The functional enrichment study was performed using Metascape. LASSO analysis was performed using the "glmnet" package in R, while the random forest algorithm was realized using the "randomForest" package in R. The prognostic variables were evaluated using univariate Cox analysis. The prognostic risk model was built using the LASSO approach. ROC curve and KM survival analyses were performed to assess the prognostic potential of the risk model. The half-maximal inhibitory concentration (IC50) of the PARP inhibitors was estimated using the "pRRophetic" package in R and the Genomics of Drug Sensitivity in Cancer database. The "rms" package in R was used to create the nomogram. A high HRD score indicated a poor prognosis and an advanced clinical process in PAAD patients. PAAD tumors with high HRD levels revealed significant T helper lymphocyte depletion, upregulated levels of cancer stem cells, and increased sensitivity to rucaparib, Olaparib, and veliparib. Using WGCNA, 11 coexpression modules were obtained. The red module and 122 hub genes were identified as the most correlated with HRD in PAAD. Functional enrichment analysis revealed that the 122 hub genes were mainly concentrated in cell cycle pathways. One novel HRD-related gene signature consisting of CKS1B, HJURP, and TPX2 were screened via LASSO analysis and a random forest algorithm, and they were validated using independent validation sets. No direct association between HRD and *CKS1B*, *HJURP*, or *TPX2* has not been reported in the literature so far. Thus, these findings indicated that *CKS1B*, *HJURP*, and *TPX2* have potential as diagnostic and prognostic biomarkers for PAAD. We constructed a novel HRD-related prognostic model that provides new insights into PAAD prognosis and immunotherapy. Based on bioinformatics analysis, we comprehensively explored the therapeutic and prognostic potential of HRD in PAAD. One novel HRD-related gene signature consisting of CKS1B, HJURP, and TPX2 were identified through the combination of WGCNA, LASSO analysis and a random forest algorithm. A novel HRD-related risk model that can predict clinical prognosis and immunotherapeutic response in PAAD patients was constructed.

**ARTICLE HISTORY**
Received 28 July 2022
Revised 3 December 2023
Accepted 5 December 2023

**KEYWORDS**
Pancreatic adenocarcinoma; tumour biomarkers; cancer immunotherapy; ICGC data portal; The Cancer Genome Atlas program

## Introduction

Pancreatic adenocarcinoma (PAAD) accounts for 90% of cancers of the pancreas and currently has a dismal 5-year survival rate of 10% [1]. Due to the insidious onset and rapid development of PAAD, the majority of patients are diagnosed with a progressive stage at the time of the initial clinical visit and lose the opportunity for surgery [2]. Thus, the exploration and development of new treatment strategies for PAAD are of utmost

**CONTACT** Huan Ouyang auyeungfoon1010@163.com Department of General surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui Province 230022, China
*These authors contributed equally: Chun Liu and Jingyun Fang.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/15384101.2023.2293594.

importance, such as immunotherapy and targeted molecular therapy.

Homologous recombination repair (HRR) is a major high-fidelity DNA damage repair (DDR) mechanism that is responsible for the slow, specialized, complicated, and precise repair of DNA double-strand breaks (DSBs) and interstrand crosslinks [3]. HRR gene mutations are common in many cancer types (13–17%), including breast, ovarian, and pancreatic cancers [4]. When HRR fails (homologous repair deficiency; HRD), such as owing to a lack of *BRCA1* or *BRCA2*, the cell must rely on alternate repair mechanisms, such as nonhomologous end-joining (NHEJ), which is less accurate and error-prone, resulting in the accumulation of new mutations and chromosomal instability, thereby increasing the likelihood of malignant transformation [5]. HRD is closely linked to genomic scarring and large-scale genome instability. Using SCNA calls obtained by ABSOLUTE, Theo A et al [6] developed an HRD score for all TCGA samples using a combination of loss of heterozygosity (LOH), large-scale state transitions (LST), and telomeric allelic imbalance (TAI). LOH was defined as the number of counts of chromosomal LOH regions shorter than whole chromosome and longer than 15 Mb. LST was defined as chromosome breakpoint (change in copy number or allelic content) between adjacent regions each of at least 10 megabases obtained after smoothing and filtering shorter than 3 Mb small-scale copy number variation. TAI was defined as the number of regions with allelic imbalance which extend to the subtelomere but do not cross the centromere. The HRD score was defined as the sum of TAI, LST, and LOH scores. Hundreds of thousands of SNV sites need to be detected to comprehensively assess the genomic scarring for the HRD score, and this process is complicated and costly. Therefore, the exploration of novel and convenient markers for HRD is highly warranted for clinical application [7].

Poly (ADP-ribose) polymerase (PARP) inhibitors, such as olaparib, have been approved by the FDA for the clinical treatment of PAAD. Encouraging therapeutic effects of *PARP* inhibitors have been observed, particularly in patients with HRD. The potential of HRD for predicting treatment responses to PARP inhibitors has been widely reported. In addition, the interaction between HRD and the ICI immunotherapy response in several cancer types has been reported [8–11]. DNA damage has been shown to increase the expression of *PD-L1* at the cell surface by promoting the expression of *PD-L1* mRNA. This regulation is dependent on the activity of *ATM-ATR/CHK1* signaling transduction, and HRD may lead to more *PD-L1* overexpression triggered by double-strand breaks (DSBs) [12]. Immunotherapy focuses on immunological checkpoints, such as *PD-1* and its ligand, *PD-L1*. The presence of *PD-L1* has been demonstrated to affect the efficacy of *PD-1/PD-L1*-targeted immunotherapies [13].

However, the potential of HRD in cancer prognosis and therapeutics has not been fully explored in PAAD. Furthermore, discovering novel biomarkers to assess HRD levels would be advantageous, and it may also increase the therapeutic potential of HRD-targeting drugs in PAAD [14].

## Materials and Methods

### *Data acquisition*

We obtained the RNA sequencing (RNA-seq) data and clinical information of PAAD patients from The Cancer Genome Atlas (TCGA) (TCGA-PAAD, $n$ = 176) [15]. In addition, four microarray data cohorts were downloaded from the Gene Expression Omnibus (GEO) (GSE49481, GSE54219, and GSE54265) and International Cancer Genome Consortium (ICGC) (ICGC-PACA-CA, $n$ = 165) as independent test sets [16,17]. Information about BRCA1 mutations in each tumor sample was provided in the GSE49481 and GSE54219 cohorts. We utilized the "IMvigor210CoreBiologies" package in R to obtain the transcriptome data and clinical information of the IMvigor 210 cohort [18].

Subsequently, the GEPIA database was utilized to acquire the differentially expressed genes (DEGs) in PAAD and normal tissues with the screening criteria set as |Log2FC|>1 and FDR< 0.01 ($n$ = 9221; Supplementary Table S1) [19]. The HRD score of each patient in TCGA-PAAD cohort was evaluated following as previously reported and the the calculation formula is as follows:HRD score = LOH+LST+TAI [20] (Supplementary Table S2). The stromal, immune, and ESTIMATE scores for each sample in TCGA-

PAAD were downloaded from the ESTIMATE database (https://bioinformatics.mdanderson.org/estimate/). The mRNA expression-based stemness index (mRNAsi) for each patient from TCGA-PAAD was obtained from a previous study [21]. Because the mRNAsi indicates the dedifferentiation potential of cancer cells, it is considered a marker for cancer stem cells (CSCs).

### Identification of immune infiltration subtype characterization of PAAD

We used the Immune Cell Abundance Identifier (ImmuCellAI) database to assess the abundance of 24 types of immune cells in TCGA-PAAD cohort by inputting RNA-seq data (Supplementary Table S3) [22]. Unsupervised hierarchical clustering of the 24 immune cell fractions was performed by the "ConsensusClusterPlus" package in R to identify the different immune infiltration subtypes in TCGA-PAAD [23].

### Weighted correlation network analysis (WGCNA)

The transcriptional profiles of the 9221 DEGs were input into the "WGCNA" package in R to construct the coexpression networks, and the detailed process is described in Supplementary Figure 1. To distinguish modules with different expression patterns, a soft threshold power was used for creating coexpression networks. Pearson's correlation analysis was performed to estimate the correlation between module eigengenes (MEs) and the HRD score, and the module with the highest Pearson's coefficient was identified as the module most relevant to the HRD score (HRD-related module). Based on the criteria of $|MM| > 0.8$ and $|GS| > 0.1$, we screened the distinct hub genes in the HRD-related module [24].

### Gene enrichment analysis

The Metascape database was utilized to perform enrichment analyses. Terms with a $p\ value < 0.01$, minimum count of 3, and an enrichment factor > 1.5 were utilized in the next step of the analysis [25]. Using screening criteria of kappa scores = 4 and similarity > 0.3, Metascape was utilized to perform hierarchical clustering to partition enrichment terms into distinct clusters, and the terms with the minimum $p\ value$ were selected as the representative terms.

### Drug sensitivity prediction

The half inhibitory concentration (IC50) of chemotherapeutic agents (AG.014699, AZD.2281, and ABT.888) was evaluated in every PAAD specimen of TCGA-PAAD by the "pRRophetic" package in R [26]. Ridge regression was performed by the "pRRophetic" algorithm, and internal cross validations were conducted by tenfold cross validation. The above predictions were performed based on the Genomics of Drug Sensitivity in Cancer (GDSC) database [27].

### Machine learning for the candidate HRD-related gene signatures (HRGSs)

In TCGA-PAAD cohort, the LASSO and the randomForest algorithms in the "glmnet" and "randomForest" packages in R, respectively, were utilized to screen candidate HRGS [28–30]. The dependent variable in all algorithms was HRD status (high versus low).The following parameters were set in the LASSO algorithm: family = "binomial"; alpha = 1; type.measure = "deviance"; and nfolds = 10. The "randomForest" package in R was used to grow a forest of 500 trees using the default settings. We computed feature importance scores with the random forest model using the "importance" function in the "random forest" package in R. Based on the randomForest algorithms, we selected the top five genes with the highest importance for downstream analysis. The intersection of the results between the two methods was identified as HRGS.

### HRGS protein subcellular localization

The Human Protein Atlas (HPA) database was utilized to visualize the relative protein distribution and subcellular localization of HRGSs in cancer cells (https://www.proteinatlas.org/) [31]. The variation in HRGS protein or transcript expression correlated with the cell cycle was obtained from the HPA database.

### Statistical analyses

R software (version 4.0.4) was utilized for all statistical procedures. Continuous variables were compared with the Wilcoxon/Kruskal–Wallis test. Differences in proportion were tested by the chi-square test. A *p value* less than 0.05 was considered significant. Kaplan–Meier (KM) survival analysis for overall survival (OS) was performed between different subgroups followed by the log-rank test. A receiver operating characteristic (ROC) curve was constructed to assess the predictive efficacy of the prognostic prediction model. For correlation analysis, Spearman's correlation was applied [32].

## Results

### Prognostic potential of the HRD score in PAAD

A total of 165 PAAD patients had an HRD score and were subsequently classified into two groups (high/low HRD group) according to the median HRD score (15). A significant difference in prognosis was observed between the high/low HRD groups (HR = 1.58; log-rank test $p = 0.031$; Figure 1a). Male sex was less frequent in the low HRD group (44% vs. 65%; chi-square test, $p < 0.05$; Figure 1b), and there were higher proportions of stage I and G1 tumors in the low HRD group (chi-square test, $p < 0.05$).

Genome instability is closely connected to immune cell infiltration in a variety of tumor types, and different immune microenvironments (TMEs) lead to differences in prognosis. Inefficient or defective homologous repair indirectly contributes to genome instability. Therefore, we further explored the differential immune infiltration pattern between the high and low HRD groups (Figure 2). Patients in TCGA-PAAD were clustered based on ImmuCellAI data by unsupervised clustering with "ConsensusClusterPlus", and three distinct immune subtypes were identified, namely, Cluster A, Cluster B, and Cluster C (Figure 2a). The high HRD group demonstrated higher proportions of the Cluster B subtype and a lower proportion of the Cluster C subtype (Figure 2b). Notably, Cluster B exhibited a distinct downregulated infiltration degree of T helper lymphocytes (Thl; CD4-T, Th1, Th2, Th17, Tfh, and n/i Treg) and B-lymphocytes as indicated by a heatmap (Figure 2c). Cluster C exhibited a distinct downregulated infiltration degree of dendritic cells (DCs), gamma delta-T cells and macrophages. In addition, the stromal, immune, and ESTIMATE scores were significantly higher in the low HRD group than in the high HRD group (Figure 2d). Moreover, the low HRD group had a markedly diminished mRNAsi score (Wilcoxon Test, $p = 0.0013$; Figure 2e).

### HRD-related gene module revealed by WGCNA

The soft threshold for network construction was set to 12 (Supplementary Figure S2A), and 11 modules, including 4516 genes, were identified (Supplementary Figure S2B). The MEs of modules were applied to calculate Pearson's correlation coefficients between the modules and HRD scores. Subsequently, we identified the red module as the most significantly related to the HRD score (Pearson's correlation $r = 0.43$, $p < 0.0001$; Supplementary Figure S2C). The association between MM and GS for the HRD score was evaluated in each module, and we screened 122 distinct hub genes in the red module based on the criteria of $|MM| > 0.8$ and $|GS| > 0.1$ (Supplementary Figure S3, Supplementary Table S4).

Functional enrichment analysis of 122 hub genes in the red module was performed using the Metascape database (Figure 3a; Supplementary Table S5), and these hub genes were mainly enriched in cell cycle-regulating pathways.

HRD is an important predictive biomarker for sensitivity to PARP inhibitors. In the GDSC database, complete drug sensitivity data of three different PARP inhibitors were available, including ruaparib (AG014699), olaparib (AZD2281), and veliparib (ABT-888). The "pRRophetic" algorithm indicated that higher IC50 values for these three PARP inhibitors were in the high HRD group (Wilcoxon test, $p < 0.001$; Figure 3b), which suggested a lower HRD score, indicating a higher sensitivity to rucaparib (AG014699), olaparib (AZD2281) and veliparib (ABT-888).

### Identification of candidate HRGSs using machine learning

The 122 hub genes of the red module were analyzed by LASSO and the random forest

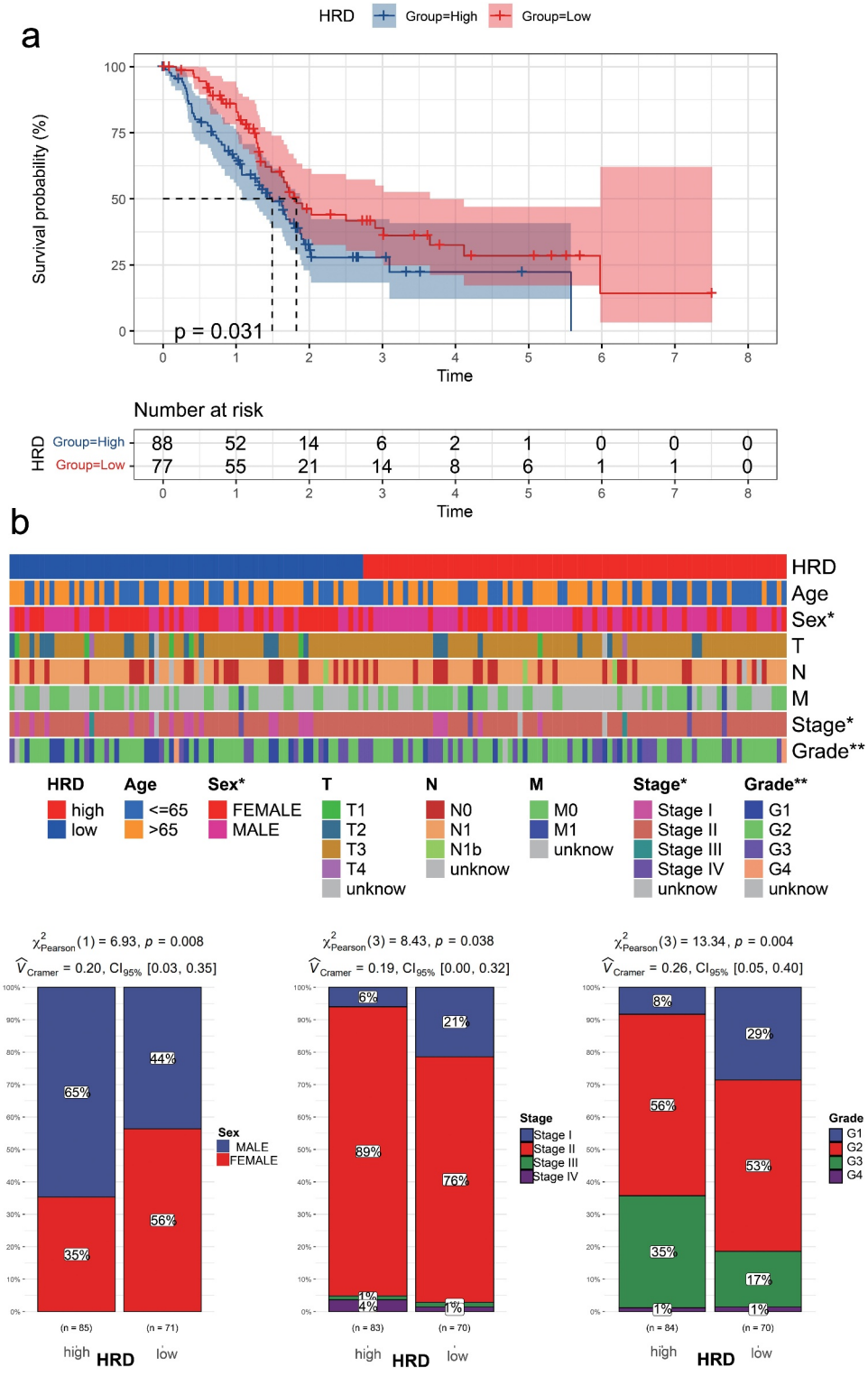


**Figure 1.** (a) The patients were classified into the high risk or low-risk group according to the median HRD score. The Kaplan – Meier curve indicated a high level of HRD score, suggesting that patients have a poor prognosis (log rank *p*-value = 0.031). (b) the proportion of PAAD patients at different sex, tumor stage and tumor grade in high/low HRD score group.

algorithm. The overlapping three genes (*CKS1B*, *HJURP*, and *TPX2*) were identified as HRGSs, and they were all highly expressed in the high HRD group (Wilcoxon Test, $p < 0.0001$; Figure 4a,b). Information about the pathology and genome functions of these three genes was extracted from reviewing relevant publications and is summarized in Supplementary Table 6.

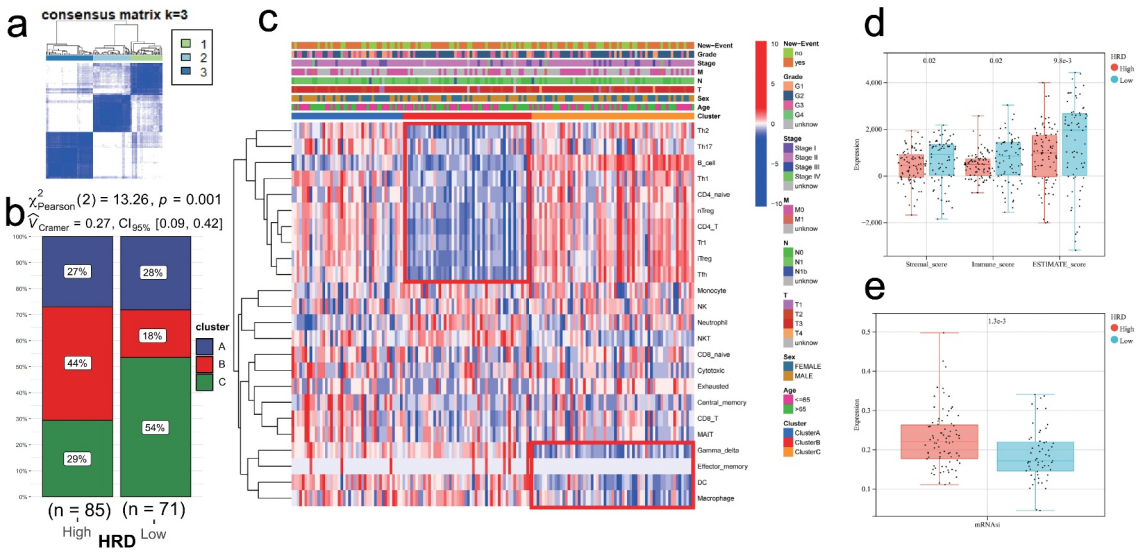


**Figure 2.** Identify different immune infiltration subtypes in TCGA-PAAD. (a) consensus matrices of the TCGA-PAAD cohort for k = 3. (b) the proportion of immune infiltration subtypes in the high/low HRDscore group, *p*-values were from the chi-squared test (cluster A, blue; cluster B, red; cluster C, green). (c) heat map comparing the abundance of 24 immune cells in TCGA-PAAD cohort among the different immune infiltration subtypes. (d) difference in the stromal, immune, and ESTIMATE scores according to high/low HRD score group. (e) difference in mRnasi level according to high/low HRD score group.

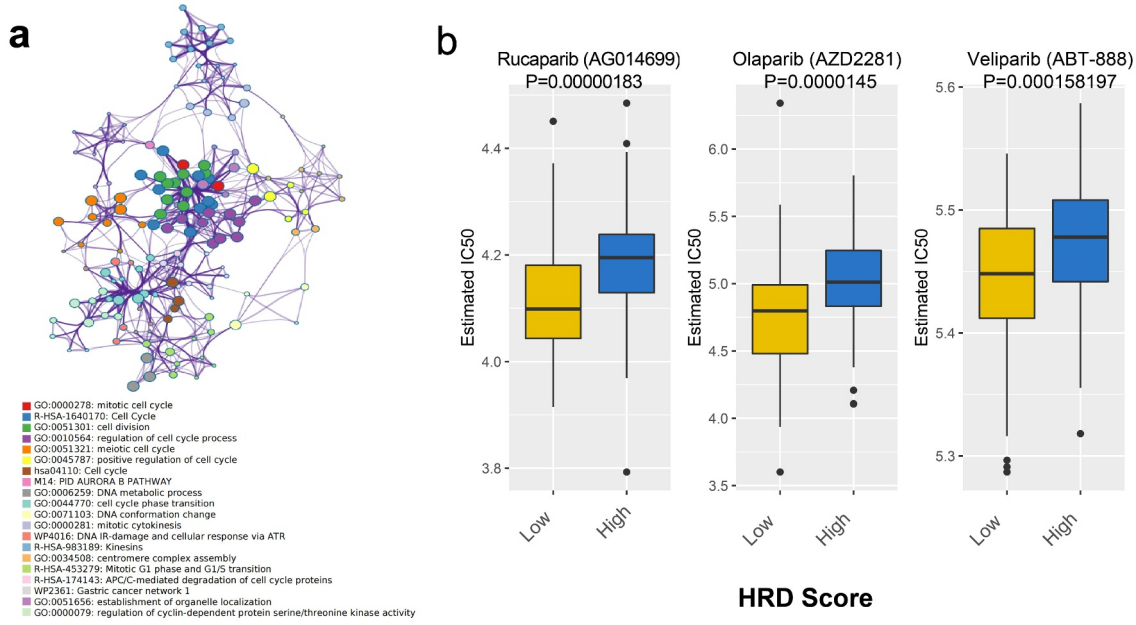


**Figure 3.** (a) Network of enriched terms of the 122 hub genes in the red module by Metascape. Colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (b) we evaluated the IC 50 of three different PARP inhibitors [rucaparib (AG014699), olaparib (AZD2281) and veliparib (ABT-888)] between high and low HRD score by performing the R package "pRrophetic".

Subsequently, the diagnostic and prognostic values of HRGSs in PAAD were further explored based on the GEPIA database. HRGSs were all highly expressed in PAAD tumors and were all unfavorable prognostic markers of PAAD patients (Figure 4c-h). Samples in TCGA-PAAD were classified into low and high expression groups according to the median expression of HRGSs. The "pRRophetic" algorithm indicated that the high expression groups exhibited

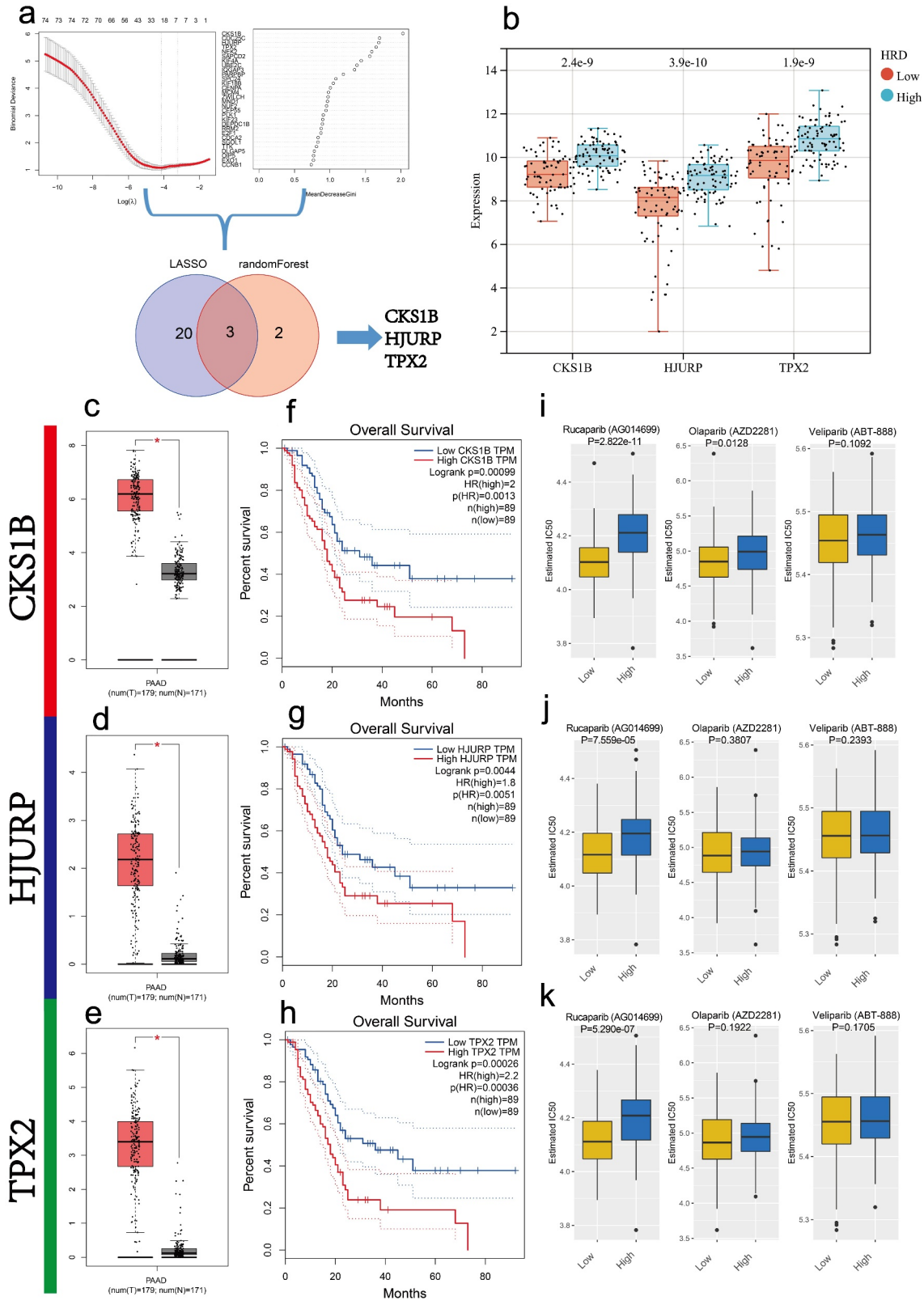


**Figure 4.** (a) Flow chart of the screening procedure. (b) variation in gene expression levels of CKS1B, HJURP and TPX2 between high and low HRD score groups (*p*-values were calculated by the Wilcoxon test). With the GEPIA database, we compared the expression levels of CKS1B (c), HJURP (d) and TPX2 (e) in PAAD tumors and normal tissues (red color represents tumor tissue, and black color represents normal tissue; red asterisk represents a *p*-value ≤0.01). Kaplan-Meier curves (OS) for expression of CKS1B (f), HJURP (g) and TPX2 (h) (*p*-values were obtained from log rank test). Differences in IC 50 of rucaparib (AG014699), olaparib (AZD2281) and veliparib (ABT-888) evaluated by R package "pRrophetic" between the high and low groups of CKS1B (i), HJURP (j) and TPX2 (k).

a higher IC50 for rucaparib, suggesting the predictive potential of HRGSs for drug sensitivity to rucaparib (Figure 4i-k).

### Subcellular localization of HRGS protein

The subcellular localization of HRGS proteins was investigated using immunofluorescence-confocal laser scanning microscopy in the HPA database. *CKS1B* was found mostly in the cytosol and mitochondria with minor levels in the nucleoplasm and vesicles (Supplementary Figure S4A). *HJURP* localized to the nucleoplasm, nucleoli and nucleoli rim, and small quantities of *HJURP* were found in mitochondria (Supplementary Figure S4B). *TPX2* was predominantly located in the nucleoplasm and to a limited extent in the cytokinetic bridge and mitotic spindle (Supplementary Figure S4C). Among different types of human cancer cell lines, there were no significant changes in the subcellular localization of target proteins (Supplementary Table S7). *CKS1B* had the highest RNA expression in the S and G2 phases (Supplementary Figure S4D). *HJURP* and *TPX2* had relatively lower protein expression in the G1 phase but the highest expression in the S and G2 phases (Supplementary Figure S4E and F).

### Validation of HRGSs as biomarkers for HRD

To determine whether *CKS1B*, *HJURP*, and *TPX2* are markers for HRD, two validations were performed. First, Spearman's analysis was performed to estimate the correlation between the HRD score and the gene expression of HRGSs based on TCGA pancancer data. In the majority of cancers, HRGSs had significant positive correlations with HRD scores ($p < 0.0001$; Figure 5a-c). Because *BRCA1* mutation is a well-known contributor to HRD, we analyzed two transcriptomics datasets from the GEO database (GSE49481 and GSE54219), each with *BRCA1* mutant cancer samples and *BRCA1* wild-type controls. The expression levels of the three HRGSs in the BRCA1-mutant group were higher than those in the *BRCA1* wild-type group ($p < 0.01$; Supplementary Figure S5A and B). ROC curves were generated to estimate the diagnostic value of these three HRGSs for *BRCA1* mutation. In the present study, the area under the curve (AUC) of all the above HRGSs was more than 0.65, and the area under the curve of *TPX2* was the highest with a value of 0.78 in the GSE49481 dataset (Supplementary Figure S5C and D).

To further validate the prognostic value of the HRGSs, the Long-term Outcome and Gene Expression Profiling Database of pancancers (LOGpc), which contains a total of 1302 PAAD samples available for prognostic information, was used. The high and low groups were classified according to the median expression levels of HRGSs. High expression of HRGSs was a risk factor for poor PAAD prognosis, which was consistent with prior observations in Figure 4 (Supplementary Figure S6A-C).

### HRD-related prognostic model of PAAD and its usefulness in immunotherapy

The LASSO algorithm identified two prognostic genes from the HRGSs (HJURP and TPX2), which were utilized to construct the HRD-related prognostic model (Figure 6a). The risk scores of every patient in TCGC-PAAD were evaluated by the HRD-related prognostic model with the following formula: risk score =*HJURP* *0.0750092401451052+*TPX2* *0.294086397668997. Patients in TCGC-PAAD were divided into low and high risk score groups according to the median risk score (40.088). The K-M survival analysis indicated that the low-risk score group had a better prognosis (HR = 1.85, 95% CI (1.22–2.82); log rank test $p = 0.0036$; Figure 6b), and the ROC curve of 5-year OS revealed a good predictive value (AUC = 0.74) (Figure 6c). To verify the robustness of the HRD-related prognostic model, used an independent dataset (ICGC-PACA-CA) for validation. A similar result was obtained in the ICGC-PACA-CA cohort (Supplementary Figure S6D). A high risk score was also a significant risk factor for poor prognosis in the ICGC-PACA cohort (HR = 1.85, 95% CI (1.34, 2.54); log rank test $p = 0.00012$). Univariate Cox analysis indicated that high risk score, age and lymph node metastasis (N) were risk factors affecting the prognosis of PAAD patients in TCGA-PAAD. Therefore, we combined age, lymph node metastasis and risk score to construct a nomogram (Figure 6d).

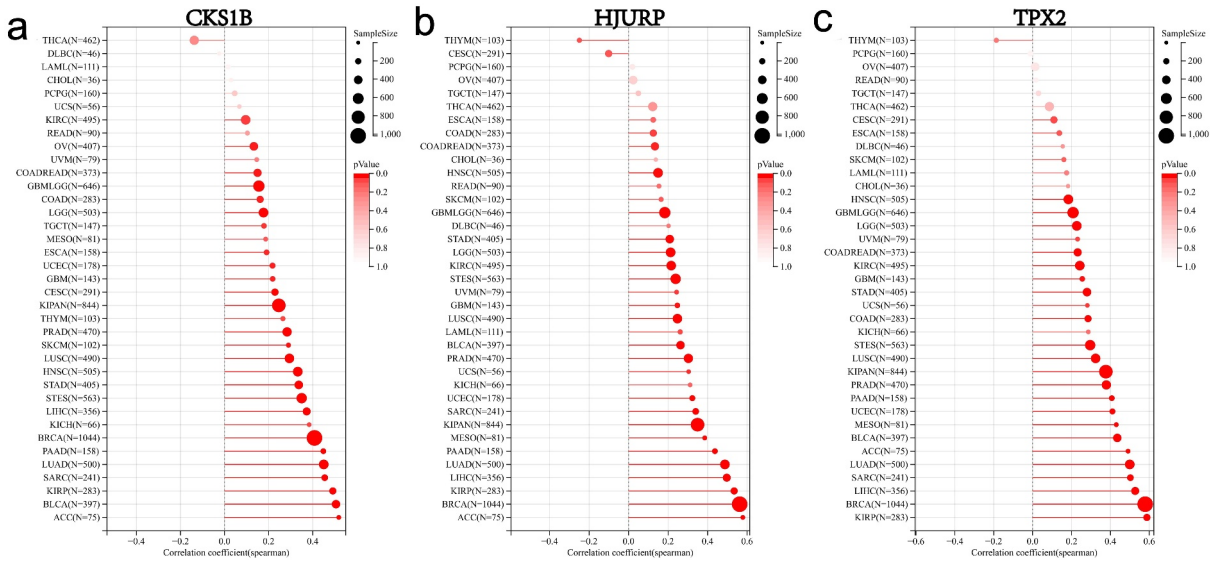


**Figure 5.** pan-cancer correlation analysis between the expression levels of CKS1B (a), HJURP (b) and TPX2 (c) and the HRD score in different kinds of tumors using Spearman analysis.

The differences in HRD scores suggested distinctly different immune infiltration in PAAD (Figure 2b). Hence, we further explored the relationship between the HRD-related prognostic model and immune infiltration in TCGA-PAAD cohort. Seven algorithms were utilized to investigate the differences in immune cell infiltration between the low- and high-risk score groups (Figure 7a). According to the heatmap, the low-risk score group had a relatively higher level of immune cell infiltration than the high-risk score group. After exclusion of an outlier (TCGA-IB-7651, tumor mutation burden (TMB) = 306.21), the TMB of the high-risk score group was significantly higher than that of the low-risk score group ($P < 0.0001$; Figure 7b). In addition, the expression level of *PDL1* was significantly higher in the high-risk score group ($P < 0.001$; Figure 7c). These results indicated that the HRD-related prognostic model was closely related to TMB and *PDL1* expression levels, both of which have been reported as biomarkers for immunotherapy. The risk scores of every patient in the IMvigor 210 cohort were evaluated by the same formula as the HRD-related prognostic model, and ROC curves were utilized to estimate the ability of the newly developed model to predict the efficacy of immunotherapy. The findings indicated that the HRD-related prognostic model can serve as a promising predictive marker for immunotherapy (AUC = 0.67; Figure 7d). Moreover, the proportion of patients with a complete remission rate (CR) or partial remission rate (PR) in the high-risk score group was higher than that in the low-risk score group (chi-square test, $P < 0.001$; Figure 7e).

According to the K-M analysis, the high-risk score group had a better prognosis in the IMvigor210 cohort compared to the low-risk score group (HR = 0.68, 95% CI (0.52–0.89); log rank test $p = 0.0044$; Figure 8a). The IMvigor210 cohort had the greatest number of patients with bladder cancer (56%); therefore, we estimated the impact of high- and low-risk scores on the prognosis of patients with only bladder cancer. The results were similar, and the high-risk score group had a better prognosis (HR = 0.70, 95% CI (0.49–1.01); log rank test $p = 0.05$; Figure 8b). We further explored whether the significant differences in prognosis are due to the predictive power of the prediction model itself or to the high response rate to immunotherapy in the high-risk score group. We calculated the risk scores of every patient in TCGA-BLCA cohort, which included 396 bladder cancer patients who had not been previously treated with immunotherapy, by the same formula as the HRD-

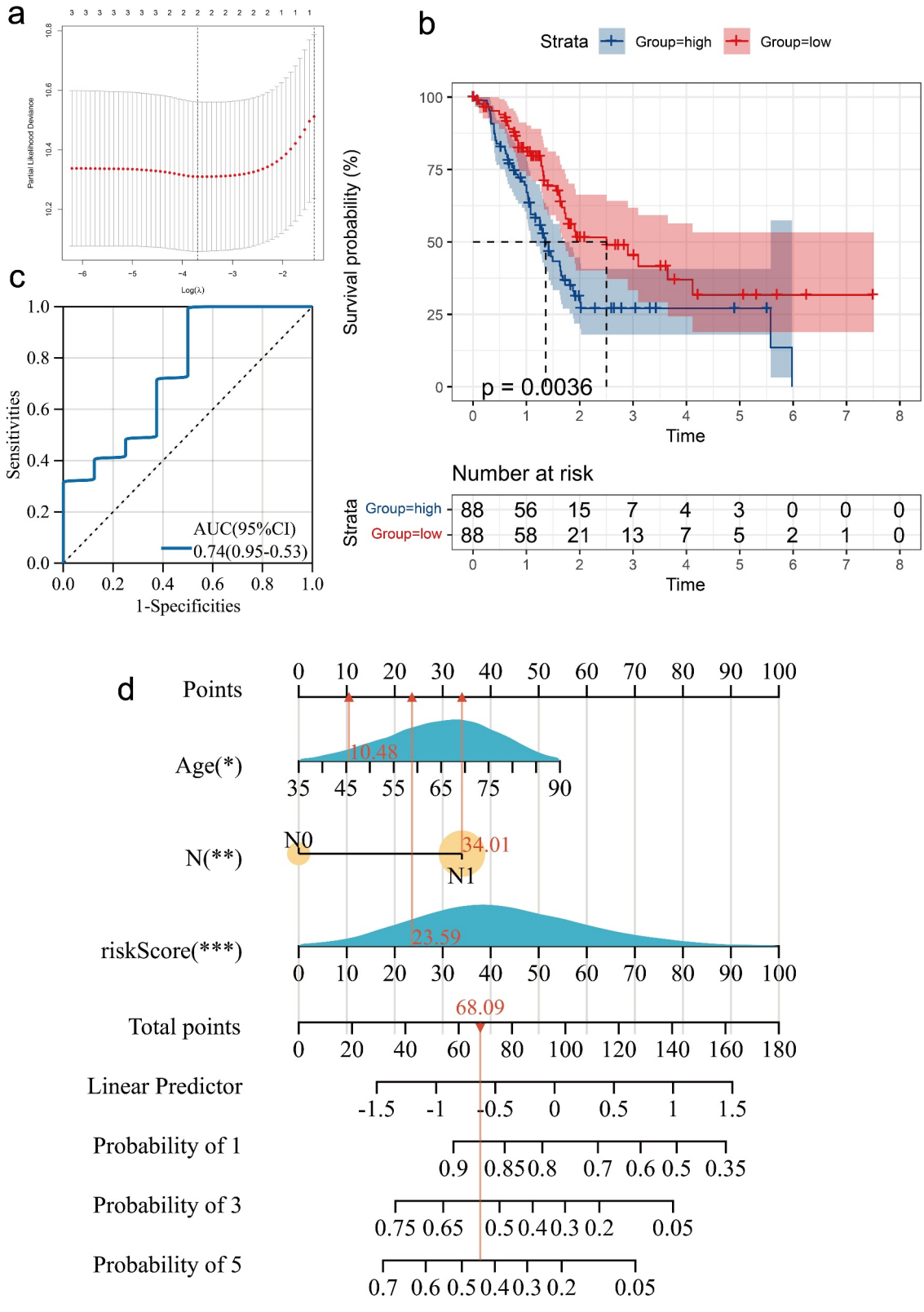


**Figure 6.** (a) risk score prognostic model construction by LASSO logistic regression analysis. (b) the patients were classified into the high risk or low-risk group according to the median risk score calculated by the prognostic model. The Kaplan – Meier curve indicated a high level of risk score, suggesting that patients have a poor prognosis (log rank *p*-value = 0.0036). (c) ROC curve of prognostic model predicting 5-year OS. (d) developed nomogram to predict the probability of survival in PAAD patients based on the age, lymph node metastasis and risk score.

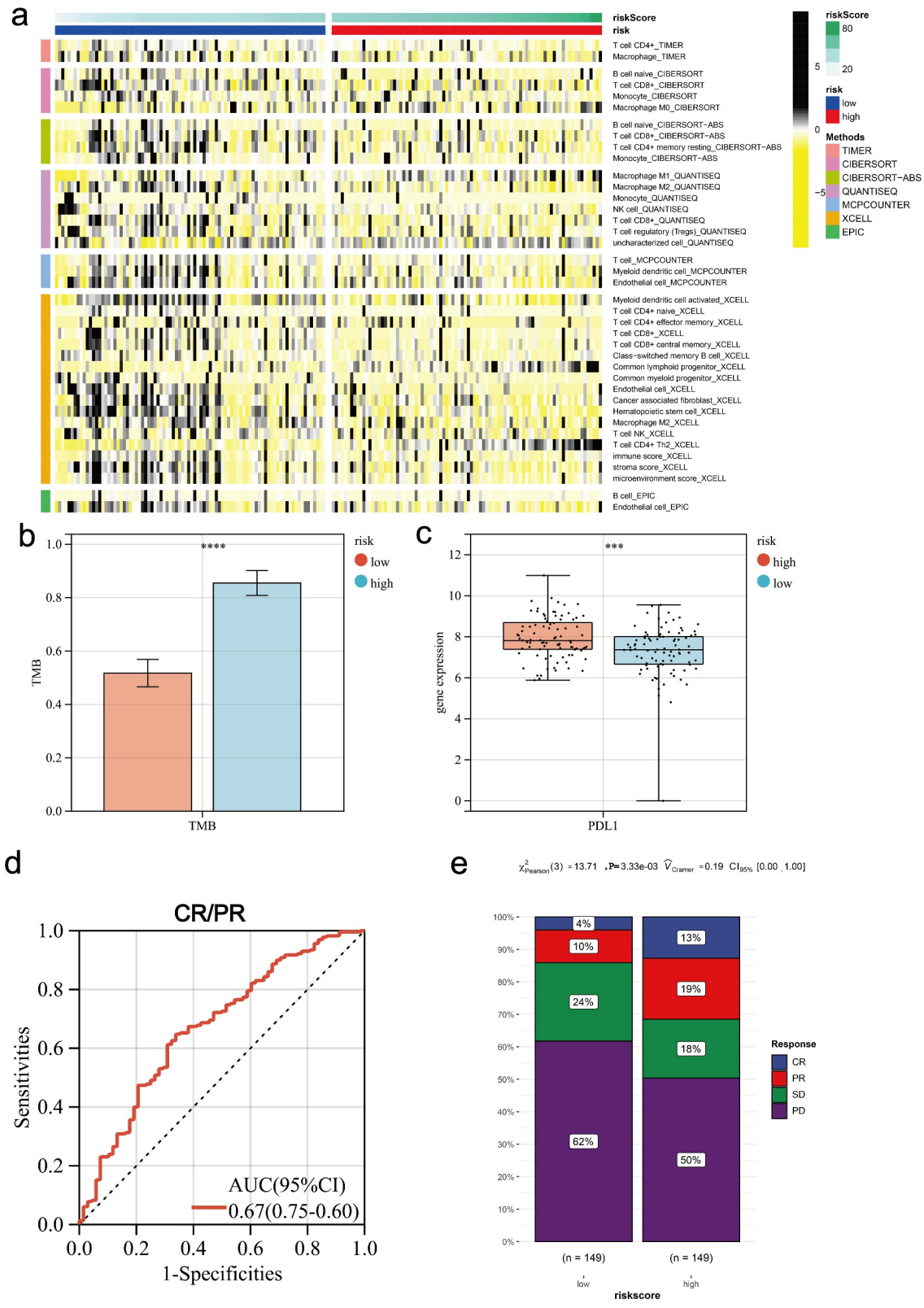


**Figure 7.** Most tumor-infiltrating immune cells showed down-regulated in high risk score group. (a) Multiple current acknowledged methods (XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT-ABS, and CIBERSORT) was used to evaluate immune cell infiltration in different risk groups. Differences in the TMB (b) and the expression levels of PD-L1 (c) between the high and low risk score groups (***$p < 0.001$;****$p < 0.0001$; $p$-values were calculated by the Wilcoxon test). (d) the predictive value of HRD-Related prognostic model in patients treated with anti-PD-1/L1 immunotherapy in IMvigor 210 cohort (AUC, 0.67). (e) the proportion of patients with response to PD-L1 blockade immunotherapy in low or high risk score groups. SD, stable disease; PD, progressive disease; CR, complete response; PR, partial response.

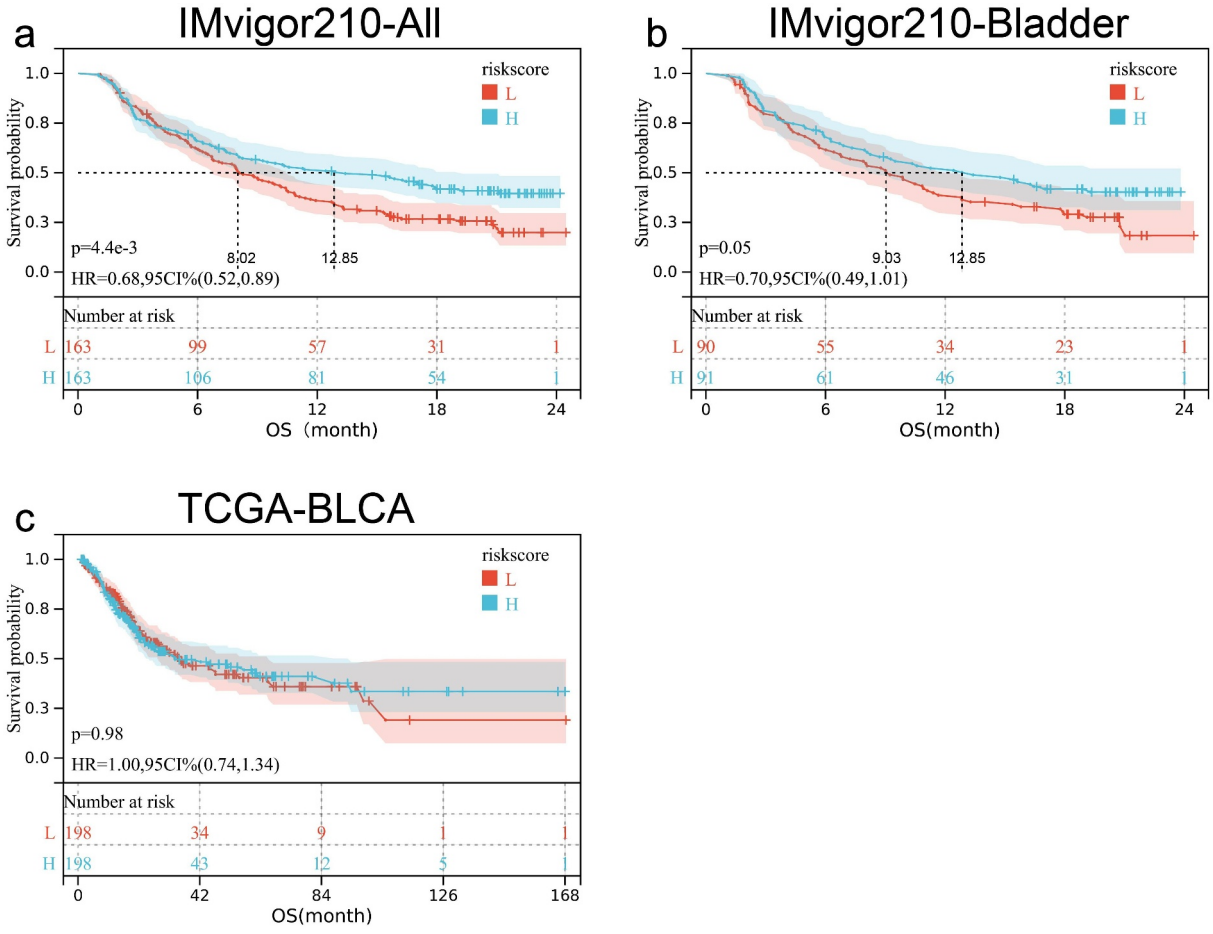


**Figure 8.** (a) Survival analyses for all low (163 cases) and high (163 cases) risk score patients in the anti-PD-L1 immunotherapy cohort using Kaplan-Meier curves (IMvigor210 cohort; HR, 0.68(0.52–0.89); $P = 0.0044$,Log-rank test). (b) survival analyses for low (90 cases) and high (91 cases) risk score patients with only bladder cancer in the anti-PD-L1 immunotherapy cohort using Kaplan-Meier curves (IMvigor210 cohort; HR, 0.70(0.49–1.01); $P = 0.05$,Log-rank test). (c) HRD-Related prognostic model can not predict patient outcome (OS) in TCGA-BLCA cohort, included 396 bladder cancer patients that had not been previously treated with immunotherapy (log rank test $p = 0.98$).

related prognostic model. However, the HRD-related prognostic model did not predict patient outcome in TCGA-BLCA cohort (log rank test $p = 0.98$; Figure 8c), which may have been due to the high response rate to immunotherapy in the high-risk score group. We validated our results using the microarray data of two additional independent immunotherapy cancer patient cohorts (GSE78220 and GSE100797). In the GSE78220 cohort, a significantly higher proportion of patients in the high-risk score group achieved CR/PR in immunotherapy (Supplementary Figure S7A). At the same time, the risk score is a good predictor of immunotherapy responsiveness (Supplementary Figure S7B). Similar to previous results, patients with high-risk scores who have a higher immunotherapy response have a better prognosis with a longer median survival (79 months *vs.* 36.6 months; log rank test $p = 0.0.03$; Supplementary Figure S7C). Similar results were also observed in the GSE100797 cohort. It is worth mentioning that all five cancer patients who achieved CR in immunotherapy were in the high-risk score group (Supplementary Figure S7D). ROC analysis suggested that the risk score has good predictive ability for immunotherapy responsiveness in the GSE100797 cohort (Supplementary Figure S7E). Additionally, patients in the high-score group also have significantly better Progression

Free Survival (PFS), which may be due to their high responsiveness to immunotherapy (79 months *vs*. 36.6 months; log rank test $p = 0.0.03$; Supplementary Figure S7F).

## Discussion

The present study comprehensively investigated the therapeutic and prognostic potential of HRD in PAAD based on bioinformatics analysis. HRD is largely responsible for genomic scarring (LOH, LST, and TAI) with large-scale genome instability. Because the HRD score was developed based on the combination of LOH, LST, and TAI to estimate the HRD status [6], the selection of the HRD score as an indicator of HRD status was reasonable for the present study. The HRD score has been extensively used and cited in numerous studies of high quality [33]. It has been reported that positive HRD status can be used as theranostic markers of response to platinum- and PARP inhibitor-based chemotherapies. Thus, precision judgment of HRD status is of high importance to assist in formulating clinical decisions for cancer patients [34]. However, using BRCA 1/2 mutation alone as a diagnostic criteria, about 40% cancer patients with HRD would be missed [35]. Utilizing more global measures, rather than relying on documented changes in specific genes, may identify more patients with HRD. The HRD score is an algorithmic assessment of three measures of tumor genomic instability (LOH, LST, and TAI) has a better diagnostic capability for HRD. However, hundreds of thousands of SNV sites need to be detected to comprehensive assess the genomic scarring for HRD score. This process is complicated and money-consuming. Therefore, the novel and convenient markers for HRD obtained in our study, is highly promising for routine application in clinical practice.

In the present study, higher HRD scores were associated with higher tumor pathology grades (Figure 1b) and CSC levels (Figure 2e). High-grade tumors are less differentiated than low-grade tumors and are more likely to develop and propagate [36]. CSCs are the underlying origins of unrestricted tumor growth and recurrence [37]. As a result, we hypothesized that high tumor pathological grades and CSCs may have a role in the poor prognosis of the high HRD score group. In addition, we detailed the immunological landscape in TCGA-PAAD, identifying unique immune infiltration patterns in groups with high and low HRD scores. It is worth noting that the high HRD score group had significantly higher levels of Thl exhaustion (Figure 2c). Previous research has indicated that Thl plays a crucial role in immune cell recruitment to the tumor microenvironment and impacts tertiary lymphoid structure (TLS) development, thereby assisting in the creation of a safe haven in the tumor mass for the generation of effective and long-lasting anticancer immune responses [38]. Thl depletion may be another key precursor of the poor prognosis in the high HRD score group because Thl has been reported as a protective factor for prognosis among multiple cancer types [39–41].

Subsequently, we utilized the HRD score to perform WGCNA and identified the HRD-related gene module, of which the genes were mainly enriched in pathways in the cell cycle, especially in the mitotic G1 phase and G1/S transition (Figure 3a). The sister chromatid is employed as a repair template in the majority of HRR instances, limiting recombination to cell cycle stages in which the sister chromatid is present, such as the S and G2 phases, and *PALB2* ubiquitylation plays a significant role in this tight regulatory mechanism [42]. A CUL3-RING ubiquitin ligase comprised of *KEAP1* in combination with CUL3-RBX1 ubiquitinates *PALB2* at the N-terminus during the G1 phase. Ubiquitination hinders *BRCA1* and *PALB2-BRCA2* from forming an association, reducing *BRCA2* recruitment and, as a result, HRR. Therefore, forced DSB repair by HRR in the G1 phase, employing a homologous chromosome as a template, may lead to negative consequences, such as LOH, which is a symptom of HRD [43]. Thus, the enrichment analysis results supported the hypotheses that the gene modules obtained by WGCNA were significantly correlated with HRD.

By combining WGCNA and a machine learning algorithm, we found three candidate HRGSs (*CKS1B, HJURP*, and *TPX2*) as biomarkers for HRD. *CKS1B* is an essential regulatory element that increases *SCF* binding to the *P27* Kip1 cyclin inhibitor and then destroys *P27* Kip1, causing the

cell to shift from G1 to S phase [44]. Furthermore, *CKS1B* is crucial not only for cell cycle control but also for transcription, DNA damage repair, cell proliferation, cell differentiation, cell senescence, cell apoptosis, protein secretion, and protein transportation [45]. As a spindle assembly factor, *TPX2* stimulates microtubule formation and promotes spindle completion during mitosis [46]. *HJURP* is a centromere-specific protein that is required for *CENP*-A centromeric localization [47], and *HJURP* suppression causes cell senescence and stops cell cycle dynamics [48]. HRGSs localize to the nucleoplasm, mitochondria, cytokinetic bridge, and mitotic spindle, which are strictly associated with the cell cycle (Supplementary Figure S4). Using the HPA database, we found that the RNA or protein expression peak phase of HRGSs emerged at the S and G2 phases when HRR is initiated. As a result, the temporal and geographical variations in HRGSs that accompany the HRR process may contribute to the potential of HRGSs as biomarkers for HRD.

Mutated *BRCA1* genes are most frequently related to HRD, and *BRCA1* proteins are required for high-fidelity DNA double-strand break repair via the HRR pathway [49]. *BRCA1* is a component of a larger genome surveillance system (BASC) [50,51]. The *MRN* complex, mismatch repair proteins (*MSH2*, *MSH6*, and *MLH1*), *BLM* syndrome helicase, and *ATM* are hypothesized to operate as sensors for DSB DNA damage [52]. *BRCA1* deficiency causes HRD, which can be treated with *PARP* inhibitors [53]. To further validate the candidate HRGSs, we investigated the relationship between HRGS expression and *BRCA1* mutation. The gene expression levels of HRGSs were significantly upregulated in *BRCA1*-mutated tumor cells, and we demonstrated that the HRGSs were indicators of *BRCA1* mutation in the GSE49481 and GSE54219 datasets (Supplementary Figure S5). Thus, the tight relationship between HRGSs and *BRCA1* mutation suggested that HRGSs are important in the HRD process.

We also constructed a risk model that included *HJURP* and *TPX2*. This model not only predicted the prognosis of PAAD patients but also the response to immunotherapy. Interestingly, K-M survival studies revealed that a high-risk score was an unfavorable prognostic factor in TCGA-PAAD cohort but a protective prognostic factor in the IMvigor 210 cohort. Patients suffering from bladder cancer comprised the demographic group with the highest percentage in the IMvigor 210 cohort. However, the model was not significantly correlated with bladder cancer prognosis in the absence of immunotherapy in TCGA-BLCA cohort (Figure 8). Thus, these opposite prediction of our risk model may be caused by the use of immunotherapy. Patients with a high-risk score achieved a higher response rate to immunotherapy in the IMvigor 210 cohort and then suppressed tumor progression, leading to a better prognosis. Such results were consistent and significant in all three independent immunotherapy cohorts (IMvigor210 cohort, GSE78220 and GSE100797) that we utilized (line 343–359). These results also suggest the dual function and application scenarios of the prognosis model that we constructed. In pancreatic cancer patients who have not undergone immunotherapy, those who have higher scores calculated by our prognosis model have a worse prognosis. However, in cancer patients who have received immunotherapy, those with higher scores have a higher probability of treatment response and a better prognosis. Such different application scenarios further emphasize the clinical significance of our prognosis model. Therefore, the present model can identify PAAD patients with poor prognosis who require immunotherapy and provide guidance for clinical treatment.

The present study provided additional information for potential alternatives for personalizing therapeutic regimens for PAAD patients. However, the present study had several limitations. Firstly, the present study focused solely on bioinformatics with no further experimental analysis based on clinical specimens. In addition, our study can only assess correlations, rather than explaining the cause-and-effect relationship between HRGSs and HRD.Furthermore, this investigation was retrospective rather than prospective. Hence, further exploration of molecular mechanisms and prospective clinical trials are warranted to validate the present results, which include the predictive potential of HRD-related risk model for the response rate of immunotherapy. However, we the present results were validated using multiple independent datasets, and the results remained credible and acceptable.[54]

## Conclusion

The present study comprehensively investigated the therapeutic and prognostic potential of HRD in PAAD based on bioinformatics analysis. PAAD patients with low HRD scores were typically more sensitive to PARP inhibitors (rucaparib, Olaparib, and veliparib) and had significantly better prognoses. In addition, the high HRD score group had significantly higher levels of Thl exhaustion. By combining WGCNA and a machine learning algorithm, we identified *CKS1B, HJURP*, and *TPX2* as potential biomarkers for HRD in PAAD. Furthermore, we constructed a two-gene risk signature and a nomogram to predict the prognosis of PAAD. Notably, the risk signature has potential for predicting the immunotherapy response. Thus, the present study contributed to the development of personalized clinical management and treatment regimens for PAAD.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research was supported by the Natural Science Foundation of Anhui Province of China (Grant No. 2108085QH308) and.Anhui University Natural Science Research Project (Grant No. KJ2021A0344).

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found in The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/), Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) and International Cancer Genome Consortium (ICGC) data portal (https://icgc.org/).

## Authors' contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

## References

[1] Ebelt ND, Zamloot V, Zuniga E, et al. Collagenase-expressing salmonella targets major collagens in pancreatic cancer leading to reductions in immunosuppressive subsets and tumor growth. Cancers (Basel). 2021;13(14):3565–3565. doi: 10.3390/cancers13143565

[2] Strohl MP, Raigani S, Ammori JB, et al. Surgery for localized pancreatic cancer: the trend is not improving. Pancreas. 2016;45(5):687–693. doi: 10.1097/MPA.0000000000000511

[3] Gonzalez D, Stenzinger A. Homologous recombination repair deficiency (HRD): from biology to clinical exploitation. Genes Chromosomes Cancer. 2021;60(5):299–302. doi: 10.1002/gcc.22939

[4] Heeke AL, Pishvaian MJ, Lynce F, et al. Prevalence of homologous recombination–related gene mutations across multiple cancer types. JCO Precis Oncol. 2018;2018(2):1–13. doi: 10.1200/PO.17.00286

[5] Ensminger M, Löbrich M. One end to rule them all: non-homologous end-joining and homologous recombination at DNA double-strand breaks. Br J Radiol. 2020;93(1115):20191054–20191054. doi: 10.1259/bjr.20191054

[6] Knijnenburg TA, Wang L, Zimmermann MT, et al. Genomic and molecular landscape of DNA damage repair deficiency across the cancer Genome Atlas. Cell Rep. 2018;23(1):239–254. doi: 10.1016/j.celrep.2018.03.076

[7] Lotan TL, Kaur HB, Salles DC, et al. Homologous recombination deficiency (HRD) score in germline BRCA2- versus ATM-altered prostate cancer. Mod Pathol. 2021;34(6):1185–1193. doi: 10.1038/s41379-020-00731-4

[8] Reardon DA, Brandes AA, Omuro A, et al. Effect of nivolumab vs Bevacizumab in patients with recurrent glioblastoma: the CheckMate 143 phase 3 randomized clinical trial. JAMA Oncol. 2020;6(7):1003–1010. doi: 10.1001/jamaoncol.2020.1024

[9] Aiello MM, Solinas C, Santoni M, et al. Excision repair crosS complementation group 1 single nucleotide polymorphisms and nivolumab in advanced non-small cell lung cancer. Front Oncol. 2020;10:1167. doi: 10.3389/fonc.2020.01167

[10] Chambon F, Osdoit S, Bagny K, et al. Dramatic response to nivolumab in xeroderma pigmentosum skin tumor. Pediatr Blood Cancer. 2018;65(2):65. doi: 10.1002/pbc.26837

[11] Kraemer KH, Tamura D, Khan SG. Pembrolizumab treatment of a patient with xeroderma pigmentosum with disseminated melanoma and multiple nonmelanoma skin cancers. Br J Dermatol. 2018;178(5):1009–1009. doi: 10.1111/bjd.16525

[12] Jiang M, Jia K, Wang L, et al. Alterations of DNA damage response pathway: biomarker and therapeutic strategy for cancer immunotherapy. Acta Pharm Sin B. 2021;11(10):2983–2994. doi: 10.1016/j.apsb.2021.01.003

[13] Feng M, Xiong G, Cao Z, et al. PD-1/PD-L1 and immunotherapy for pancreatic cancer. Cancer Lett. 2017;407:57–65. doi: 10.1016/j.canlet.2017.08.006
[14] Hoppe MM, Sundar R, Tan DSP, et al. Biomarkers for homologous recombination deficiency in cancer. J Natl Cancer Inst. 2018;110(7):704–713. doi: 10.1093/jnci/djy085
[15] Wang Z, Jensen MA, Zenklusen JC. A practical guide to the cancer Genome Atlas (TCGA). Methods Mol Biol. 2016;1418:111–141.
[16] Clough E, Barrett T. The gene expression Omnibus database. Methods Mol Biol. 2016;1418:93–110.
[17] Aaltonen LA, Abascal F, Abeshouse A. ICGC/TCGA pan-cancer analysis of whole genomes consortium. Pan-cancer analysis of whole genomes. Nature. 2020;578(7793):82–93. doi: 10.1038/s41586-020-1969-6
[18] Vander Velde N, Guerin A, Ionescu-Ittu R, et al. Comparative effectiveness of non-cisplatin first-line therapies for metastatic urothelial carcinoma: phase 2 IMvigor210 study versus US patients treated in the Veterans Health Administration. Eur Urol Oncol. 2019;2(1):12–20. doi: 10.1016/j.euo.2018.07.003
[19] Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1):W98–W102. doi: 10.1093/nar/gkx247
[20] Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4(1):2612–2612. doi: 10.1038/ncomms3612
[21] Malta TM, Sokolov A, Gentles AJ, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. Cell. 2018;173(2):338–354. doi: 10.1016/j.cell.2018.03.034
[22] Miao YR, Zhang Q, Lei Q, et al. ImmuCellAI: a unique method for comprehensive T-Cell subsets abundance prediction and its application in cancer immunotherapy. Adv Sci. 2020;7(7):1902880. doi: 10.1002/advs.201902880
[23] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26(12):1572–1573. doi: 10.1093/bioinformatics/btq170
[24] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. 2008;9(1):559–559. doi: 10.1186/1471-2105-9-559
[25] Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523–1523. doi: 10.1038/s41467-019-09234-6
[26] Geeleher P, Cox N, Huang RS. pRrophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One. 2014;9:e107468. doi: 10.1371/journal.pone.0107468
[27] Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(D1):D955–961. doi: 10.1093/nar/gks1111
[28] Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics. 2019;11(1):123–123. doi: 10.1186/s13148-019-0730-1
[29] Hanko M, Grendár M, Snopko P, et al. Random forest-based prediction of outcome and mortality in patients with traumatic brain injury undergoing primary decompressive craniectomy. World Neurosurg. 2021;148:e450–e458. doi: 10.1016/j.wneu.2021.01.002
[30] Cheung-Lee WL, Link AJ. Genome mining for lasso peptides: past, present, and future. J Ind Microbiol Biotechnol. 2019;46(9–10):1371–1379. doi: 10.1007/s10295-019-02197-z
[31] Lei Y, Yu T, Li C, et al. Expression of CAMK1 and its association with clinicopathologic characteristics in pancreatic cancer. J Cell Mol Med. 2021;25(2):1198–1206. doi: 10.1111/jcmm.16188
[32] Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18:91–93. doi: 10.1016/j.tjem.2018.08.001
[33] Telli ML, Timms KM, Reid J, et al. Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with Triple-negative breast cancer. Clin Cancer Res. 2016;22(15):3764–3773. doi: 10.1158/1078-0432.CCR-15-2477
[34] Gou R, Dong H, Lin B. Application and reflection of genomic scar assays in evaluating the efficacy of platinum salts and PARP inhibitors in cancer therapy. Life Sci. 2020;261:118434. doi: 10.1016/j.lfs.2020.118434
[35] Sharma P, Barlow WE, Godwin AK, et al. Impact of homologous recombination deficiency biomarkers on outcomes in patients with triple-negative breast cancer treated with adjuvant doxorubicin and cyclophosphamide (SWOG S9313). Ann Oncol. 2018;29(3):654–660. doi: 10.1093/annonc/mdx821
[36] Pilehchian Langroudi M, Nikbakhsh N, Samadani AA, et al. FAT4 hypermethylation and grade dependent downregulation in gastric adenocarcinoma. J Cell Commun Signal. 2017;11(1):69–75. doi: 10.1007/s12079-016-0355-5
[37] Huang Y, Zhou B, Luo H, et al. ZnAs@SiO2nanoparticles as a potential anti-tumor drug for targeting stemness and epithelial-mesenchymal transition in hepatocellular carcinoma via SHP-1/JAK2/STAT3 signaling. Theranostics. 2019;9(15):4391–4408. doi: 10.7150/thno.32462
[38] Gu-Trantien C, Loi S, Garaud S, et al. $CD4^+$ follicular helper T cell infiltration predicts breast cancer survival. J Clin Invest. 2013;123(7):2873–2892. doi: 10.1172/JCI67428
[39] Kuczek DE, Larsen AMH, Thorseth ML, et al. Collagen density regulates the activity of tumor-infiltrating T cells. J Immunother Cancer. 2019;7(1):68–68. doi: 10.1186/s40425-019-0556-6
[40] Chen Z, Wu T, Yan Z, et al. Identification and validation of an 11-Ferroptosis related gene signature and its

correlation with immune checkpoint molecules in Glioma. Front Cell Dev Biol. 2021;9:652599. doi: 10.3389/fcell.2021.652599

[41] Chen X, Jin Y, Gong L, et al. Bioinformatics analysis finds immune gene markers related to the prognosis of bladder cancer. Front Genet. 2020;11:607–607. doi: 10.3389/fgene.2020.00607

[42] Orthwein A, Noordermeer SM, Wilson MD, et al. A mechanism for the suppression of homologous recombination in G1 cells. Nature. 2015;528(7582):422–426. doi: 10.1038/nature16142

[43] Zhao X, Wei C, Li J, et al. Cell cycle-dependent control of homologous recombination. Acta Biochim Biophys Sin (Shanghai). 2017;49(8):655–668. doi: 10.1093/abbs/gmx055

[44] Jia Y, Tian Q, Yang K, et al. A pan-cancer analysis of clinical prognosis and immune infiltration of CKS1B in human tumors. Biomed Res Int. 2021;2021:1–25. doi: 10.1155/2021/5862941

[45] Shrestha S, Yang CD, Hong HC, et al. Integrated MicroRNA–mRNA analysis reveals miR-204 inhibits cell proliferation in gastric cancer by targeting CKS1B, CXCL1 and GPRC5A. Int J Mol Sci. 2017;19(1):87–87. doi: 10.3390/ijms19010087

[46] He Y, Peng L, Li J, et al. TPX2 deficiency leads to spindle abnormity and meiotic impairment in porcine oocytes. Theriogenology. 2022;187:164–172. doi: 10.1016/j.theriogenology.2022.04.031

[47] Bassett EA, DeNizio J, Barnhart-Dailey MC, et al. HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. Dev Cell. 2012;22(4):749–762. doi: 10.1016/j.devcel.2012.02.001

[48] Serafim RB, Cardoso C, Di Cristofaro LFM, et al. HJURP knockdown disrupts clonogenic capacity and increases radiation-induced cell death of glioblastoma cells. Cancer Gene Ther. 2020;27(5):319–329. doi: 10.1038/s41417-019-0103-0

[49] Nguyen LWM, Martens J, Van HA, et al. Pan-cancer landscape of homologous recombination deficiency. Nat Commun. 2020;11(1):5584–5584. doi: 10.1038/s41467-020-19406-4

[50] Babushkina NP, Postrigan AE, Kucher AN. Involvement of variants in the genes encoding BRCA1-associated Genome surveillance complex (BASC) in the development of human common diseases. Mol Biol (Mosk). 2021;55(2):278–296. doi: 10.1134/S0026893321020047

[51] Wang Y, Cortez D, Yazdi P, et al. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev. 2000;14(8):927–939. doi: 10.1101/gad.14.8.927

[52] Jiang M, Jia K, Wang L, et al. Alterations of DNA damage repair in cancer: from mechanisms to applications. Ann Transl Med. 2020;8(24):1685–1685. doi: 10.21037/atm-20-2920

[53] Jiang X, Li X, Li W, et al. PARP inhibitors in ovarian cancer: sensitivity prediction and resistance mechanisms. J Cell Mol Med. 2019;23(4):2303–2313. doi: 10.1111/jcmm.14133

[54] Oliver GR, Hart SN, Klee EW. Bioinformatics for clinical next generation sequencing. Clin Chem. 2015;61(1):124–135. doi: 10.1373/clinchem.2014.224360