Heliyon 10 (2024) e30941

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Epigenetic-related gene-based prognostic model construction and validation in prostate adenocarcinoma

Youyou Li [a,1], Chao Li [a,1], Longxiang Wu [b], Jiaren Li [a], Yu Gan [b], Shuo Tan [a], Lei Zhou [a], Wei Xiong [a], Liang Zhou [a], Cheng Li [a], Jiahao Liu [a], Dingwen Liu [a], Yichuan Wang [a], Yunlong Fu [a], Kun Yao [a], Long Wang [a,*]

[a] Department of Urology, The Third Xiangya Hospital, Central South University, Changsha, Hunan, 410013, China
[b] Department of Urology, The Xiangya Hospital, Central South University, Changsha, Hunan, 410008, China

ARTICLE INFO

Keywords:
Prostate adenocarcinoma
Epigenetics
Immunotherapy
Prognostic model
Bioinformatics

ABSTRACT

Prostate adenocarcinoma (PRAD), driven by both genetic and epigenetic factors, is a common malignancy that affects men worldwide. We aimed to identify and characterize differentially expressed epigenetic-related genes (ERGs) in PRAD and investigate their potential roles in disease progression and prognosis. We used PRAD samples from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) to identify prognosis-associated ERGs. Thirteen ERGs with two distinct expression profiles were identified through consensus clustering. Gene set variation analysis highlighted differences in pathway activities, particularly in the Hedgehog and Notch pathways. Higher epigenetic scores correlated with favorable prognosis and improved immunotherapeutic response. Experimental validation underscored the importance of CBX3 and KAT2A, suggesting their pivotal roles in PRAD. This study provides crucial insights into the epigenetic scoring approach and presents a promising prognostic tool, with CBX3 and KAT2A as key players. These findings pave the way for targeted and personalized interventions for the treatment of PRAD.

1. Introduction

Prostate cancer (PCa) ranks as the second most frequently diagnosed malignancy and the fifth leading contributor to cancer-related deaths among men globally [1,2]. Currently, diverse treatment modalities are available for PCa [3]. For localized PCa, the conventional approach involves radical prostatectomy. Nevertheless, a substantial subset of patients with PCa, approximately 25–30 %, who undergo radical prostatectomy experience progression to advanced disease stages characterized by increased recurrence rates and an unfavorable prognosis within a decade [4]. This underscores the pressing need for a comprehensive understanding of the intricate etiological mechanisms governing the initiation and progression of PCa.

Epigenetics involves the exploration of heritable and reversible patterns of gene expression that manifest independent of DNA sequence [5]. Epigenetic influences can affect the expression of oncogenes or tumor suppressor genes, ultimately fostering tumorigenesis [6]. Evaluation of the expression levels of epigenetic factors has the potential to refine prognostic precision and facilitate

* Corresponding author.
E-mail address: wanglong@csu.edu.cn (L. Wang).
[1] Youyou Li and Chao Li contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e30941
Received 7 February 2024; Received in revised form 24 April 2024; Accepted 8 May 2024
Available online 9 May 2024

tailored patient-specific interventions. Thus, a comprehensive exploration using extensive datasets is warranted to establish a broader range of epigenetic biomarkers suitable for clinical applications. However, previous studies had some notable limitations, including relatively small sample sizes and potential biases inherent in the retrospective design; further large-scale prospective studies are needed to enhance statistical power and explore interactions between factors [7]. Additionally, while microarray-based bioinformatics analysis offers insights into the molecular mechanisms and potential biomarkers for PCa, further experimental validation of epigenetic-related genes (ERGs) at multiple levels is warranted [8,9].

In this study, comprehensive analysis of prostate adenocarcinoma (PRAD) samples from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases identified 13 ERGs that were strongly associated with PRAD prognosis. Application of consensus clustering to TCGA-PRAD data revealed two discrete clusters. In addition, gene set variation analysis (GSVA) revealed differential activity levels in pathways such as Hedgehog and Notch within these clusters. Based on these 13 key genes, we used an epigenetic scoring approach for PRAD from TCGA and found that patients with higher scores had a more favorable prognosis and demonstrated increased immunotherapeutic response. We aimed not only to identify key genes that play a role in PRAD, but more importantly, we developed a prognostic model based on ERG expression levels to provide more accurate survival prediction for patients with PRAD, aiming to improve the clinical decision-making process and provide a more personalized basis for patient treatment strategies. In addition, by analyzing the role of these genes in the development and treatment response of PRAD, our findings may help identify new treatment targets and promote innovative research in PRAD treatment.

2. Materials and methods

2.1. Data acquisition

The data set is preferably from Homo sapiens, including expression profile data from PCa patients. The dataset should contain disease and control groups and the number should be greater than 10; data with clear grouping information and not containing gene knockout or drug intervention are preferred. In addition, samples with too many negative expression matrix values will be excluded. RNA sequencing data, including 553 cases (52 paracancer tissues and 501 tumor samples) and the corresponding clinical information, were sourced from the UCSC Xena database (http://genome.ucsc.edu) and standardized into fragments per kilobase per million data format [10]. Microarray data, including GSE46602 (503 tumor tissues) [11], GSE69223 (14 paracancer tissues and 36 tumor tissues) [12], GSE6919 (81 paracancer tissues and 65 tumor tissues) [13] and GSE141551 (15 paracancer tissues and 15 tumor tissues) [14], were downloaded from the GEO database. All the samples were included in the analysis.

We systematically searched all ERGs in the GeneCards database using the term epigenetic and obtained a total of 391 related genes with a relevance score >1.5 as the filtering criterion in GeneCards, and 501 relevant genes were obtained from six reference gene sets from the Molecular Signatures Database [15]. Furthermore, we searched PubMed and obtained 720 related genes [16], Eventually, a total of 57 epigenetic genes were obtained by taking the intersection of the relevant genes from these three sources (Supplementary Table 1).

2.2. Identification of differentially expressed ERGs

Gene expression data (TCGA-PRAD, GSE6919, GSE46602, GSE69223, and GSE141551) were processed using the limma package in R, including background correction and quantile normalization. Microarray data were processed using the sva package to remove batch effect. Principal component analysis (PCA) was used to verify the reproducibility of the data, and the R package ggplot2 was used to construct PCA plots. Differentially expressed genes between PRAD and paracancer tissues with $|\log2 \text{ (fold change)}| > 0$ and adjusted $P < 0.05$ were identified using limma. The identified upregulated and downregulated differentially expressed genes were individually overlapped with the ERGs, and this intersection was visually represented using Venn diagrams. The amalgamation of these intersections constituted a pool of differentially expressed genes specific to PRAD. Subsequently, distinct comparative plots were constructed for both TCGA-PRAD and GEO-PRAD datasets.

2.3. Enrichment analyses

Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed using the R package clusterProfiler [17], and the filtering criteria for significance were set as P.adj <0.05 and false discovery rate (FDR) (q.value) < 0.2. The Benjamini–Hochberg method was used for P-value correction. Moreover, in TCGA-PRAD datasets, we used c2.cp.v7.2. symbols.gmt for gene set enrichment analysis (GSEA) to identify the different molecular mechanisms between tumors and normal tissues. The parameters used for the GSEA were as follows: seed number, 2022; number of permutations, 10,000; and each gene set was required to contain a minimum of five genes and a maximum of 500 genes. The Benjamini–Hochberg method was used for P-value correction, and the criteria for identifying significantly enriched gene sets were set as P.adj <0.05 and FDR (q.value) < 0.25. GSVA was performed using h.all.v7.5.1. symbols.gmt in the R package GSVA [18].

2.4. Construction of ERG signature

The prognostic model was established to minimize overfitting using the LASSO-penalized Cox regression analysis in the glmnet R package [19]. Eventually, four genes and their coefficients were retained, and the minimum criteria determined the penalty parameter

(λ). The risk score was obtained using the formula:

$$risk\ score = (0.0171452171253261 \times DOT1L) + (0.543761947731939 \times EZH2) + (0.111609967092668 \times KAT2A)\cdots + (0.389947150504332 \times MTA1)$$

where β represents regression coefficient. The patients with PRAD were divided into high- and low-risk groups based on the median risk score. The survival analysis between the two risk groups was performed using the survminer R package, and the receiver operating characteristic (ROC) curve analysis was performed via the survival and timeROC R packages.

2.5. Consensus clustering analysis of ERGs

Consensus clustering was applied to explore the connections between ERG expression and PRAD subtypes using the k-means method. The R package ConsensusClusterPlus was used to determine the number of clusters; the clustering was performed with a range of cluster numbers from 2 to 8. The total dataset was repeatedly sampled 1000 times at 80 % of its size. The clusterAlg parameter was set to km, and the distance parameter was set to euclidean [20]. The ROC curve was plotted using the pROC R package to calculate the area under the curve (AUC) value, which was then used to evaluate the predictive performance of the 13 ERGs for PRAD clusters. The correlations between each cluster and progression-free interval (PFI) were analyzed using the survival package. The results were presented using Kaplan–Meier graphs.

2.6. Construction of a predictive nomogram

To explore the clinical prognostic significance of the identified ERGs in PRAD, we conducted univariate Cox regression analysis using the expression profiles of the selected ERGs from TCGA-PRAD dataset. Statistically significant ($P < 0.05$) genes were subsequently chosen for inclusion in a multivariate Cox regression analysis. The analysis involved building a multivariate Cox regression model and the findings were presented visually using a forest plot. Based on the outcomes of multivariate Cox regression analysis, we constructed a nomogram to predict the probabilities of disease progression-free survival at 1-, 3-, and 5-year intervals for patients with PRAD. Finally, we used the rms R package to construct nomogram plots and calibration curve plots. Decision curve analysis (DCA) is a straightforward method used to assess clinical predictive models, diagnostic tests, and molecular markers. To evaluate the predictive performance of the nomogram model for the 1-, 3-, and 5-year survival outcomes of patients with PRAD, we used the R package ggDCA to generate DCA plots. These plots assess the effectiveness of the nomogram model in predicting survival outcomes for specified time intervals.

2.7. Single sample gene set enrichment analysis (ssGSEA)

The PRAD samples were divided into high and low PRAD groups for further analysis based on their median ERGs scores. Kaplan–Meier survival curves were used to examine the effect of the signature on survival. ssGSEA was used to generate an enriched fraction of 28 immune cells in each sample.

2.8. Tumor immune dysfunction and exclusion (TIDE)

The TIDE is a comprehensive score for tumor immune dysfunction and immune escape. TCGA-PRAD data and the corresponding clinical information were estimated using the TIDE algorithm to predict the potential immune checkpoint blockade response. Higher scores indicate poorer function [21].

2.9. Cell culture

Human PCa cell lines PC-3 and DU-145 and the benign human prostate epithelial cell line BPH-1 were provided by Procell (Wuhan, China). The PC-3, BPH-1, and DU-145 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibico, USA) or 1640 (BI, Israel) supplemented with 1 % streptomycin/penicillin and 10 % FBS (BI) under saturated humidity, 37 °C, and 5 % CO2.

2.10. Real-time quantitative polymerase chain reaction (RT-qPCR)

RNA was extracted using a TRIzol kit (TaKaRa, Japan), and cDNA was synthesized using a reverse transcription kit (Novoprotein, China). RT-qPCR was performed using the SYBR Green Mix (Bimake, USA). The primer sequences are listed in Supplementary Table S2. Gene expression levels were quantified using the $2^{-\Delta\Delta CT}$ method.

2.11. Western blotting

For total protein extraction, cells were lysed using radioimmunoprecipitation assay (RIPA) buffer supplemented with 1 mmol/L phenylmethylsulfonyl fluoride (PMSF) for 15 min on ice and centrifuged at 12,000 rpm for 10 min. The supernatants were collected. Total protein quality was measured using the BCA Protein Assay Kit (Elabscience, China) according to the manufacturer's instructions.

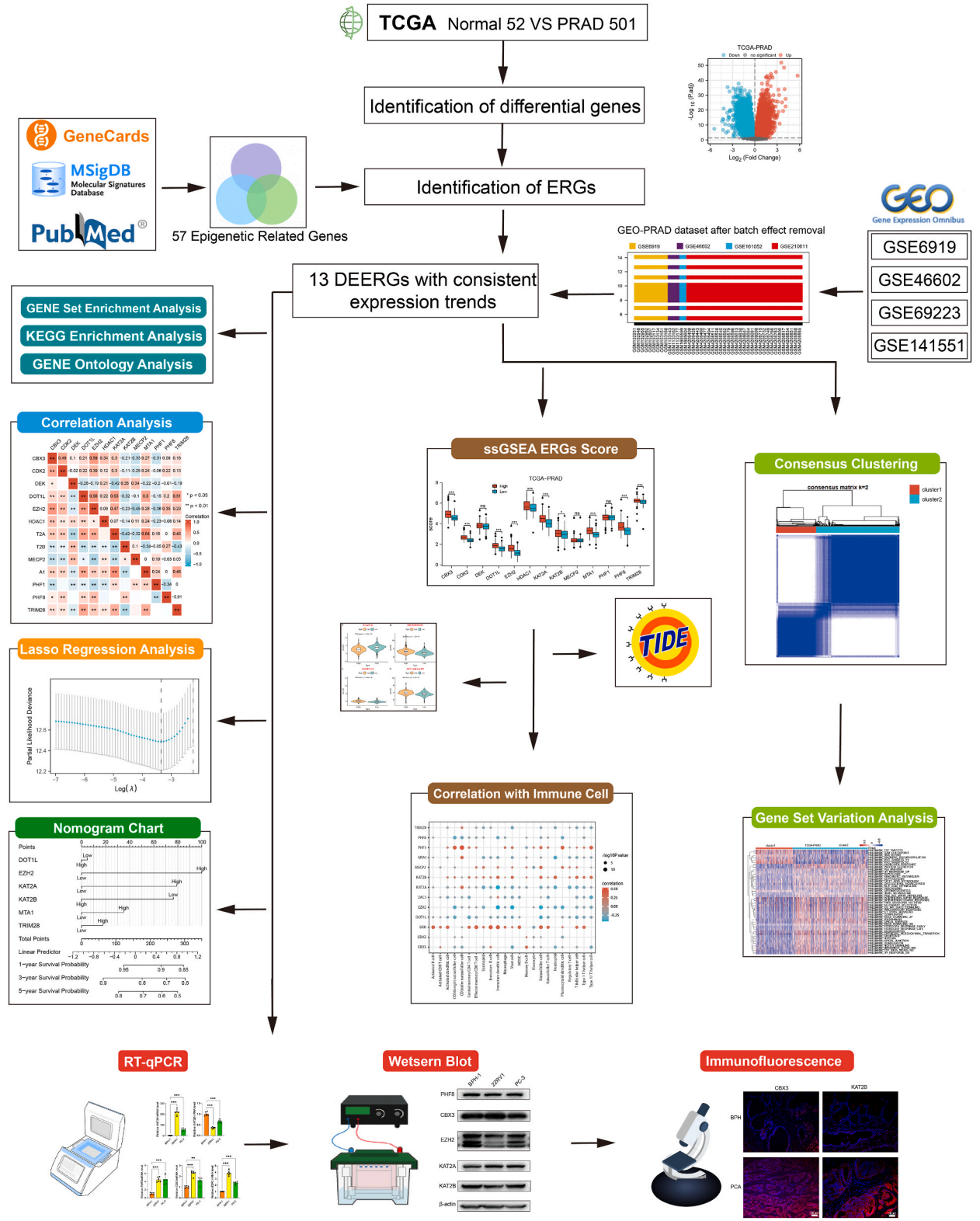


Fig. 1. A flow chart of this study.

The collected proteins were boiled for 5 min and separated using 10 % sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The separated proteins were transferred onto polyvinylidene fluoride membranes. After blocking with Quick Blocker (Beyotime, China) for 10 min at 25 °C, the membranes were incubated with specific primary antibodies (Supplemental Table 3) overnight at 4 °C and washed three times. The membranes were then incubated with horseradish peroxidase-conjugated secondary antibodies for 1 h at room temperature. The binding of each antibody was visualized using enhanced chemiluminescence (ECL), according to the manufacturer's instructions. ImageJ software was used to quantify protein expression. The steps to use the software were as follows: after saving the original image, it was opened in the ImageJ software, the quantification areas for bands were defined, and they were quantified along with the nearby background. The data were transferred to Excel and the background values were subtracted. The loading-control

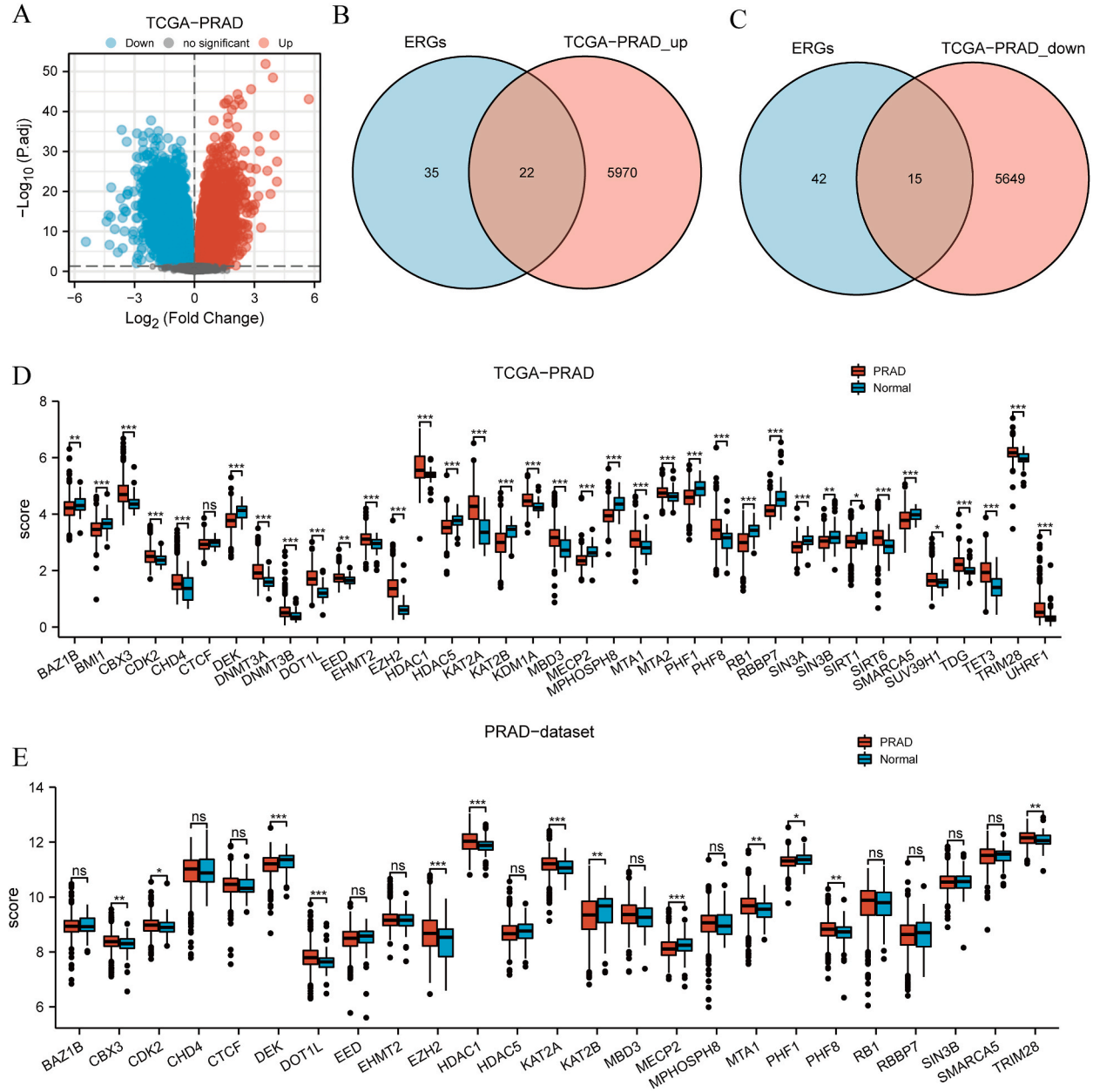


Fig. 2. Differential expression of ERGs in the TCGA-PRAD dataset and GEO-PRAD dataset.
A. The volcano plot illustrating the differential expression analysis between the PRAD group and the normal group.
B–C. The Venn diagram displaying the intersection of upregulated and downregulated differentially expressed genes obtained from the TCGA-PRAD dataset and ERGs.
D-E. Group comparison plots showing the expression levels of ERGs in the TCGA-PRAD dataset and the GEO-PRAD dataset. * indicate $P < 0.05$, ** indicate $P < 0.01$, *** indicate $P < 0.001$.

process was repeated. Normalized integrated density (IntDen) was calculated by dividing the IntDen of the protein by that of the control. The mean value of the control group was determined for the fold-change calculation. Finally, the data were statistically analyzed and visualized using GraphPad Prism.

2.12. Immunofluorescence staining

Immunofluorescence staining was performed as previously described [22]. For paraffin-embedded human PRAD or paracancer samples, 5-μm sections were deparaffinized with xylene and ethanol. Antigen retrieval was performed by microwave irradiation in citrate buffer (pH 6.0) for 20 min. Sections were then incubated with primary antibodies at 4 °C overnight. Anti-rabbit-555 secondary antibody was used. Nuclei were stained with DAPI (#H-1200-10; Vector Labs, USA). Images were captured using a KF-FL-400 microscope (KFBIO, China) using a × 40 objective.

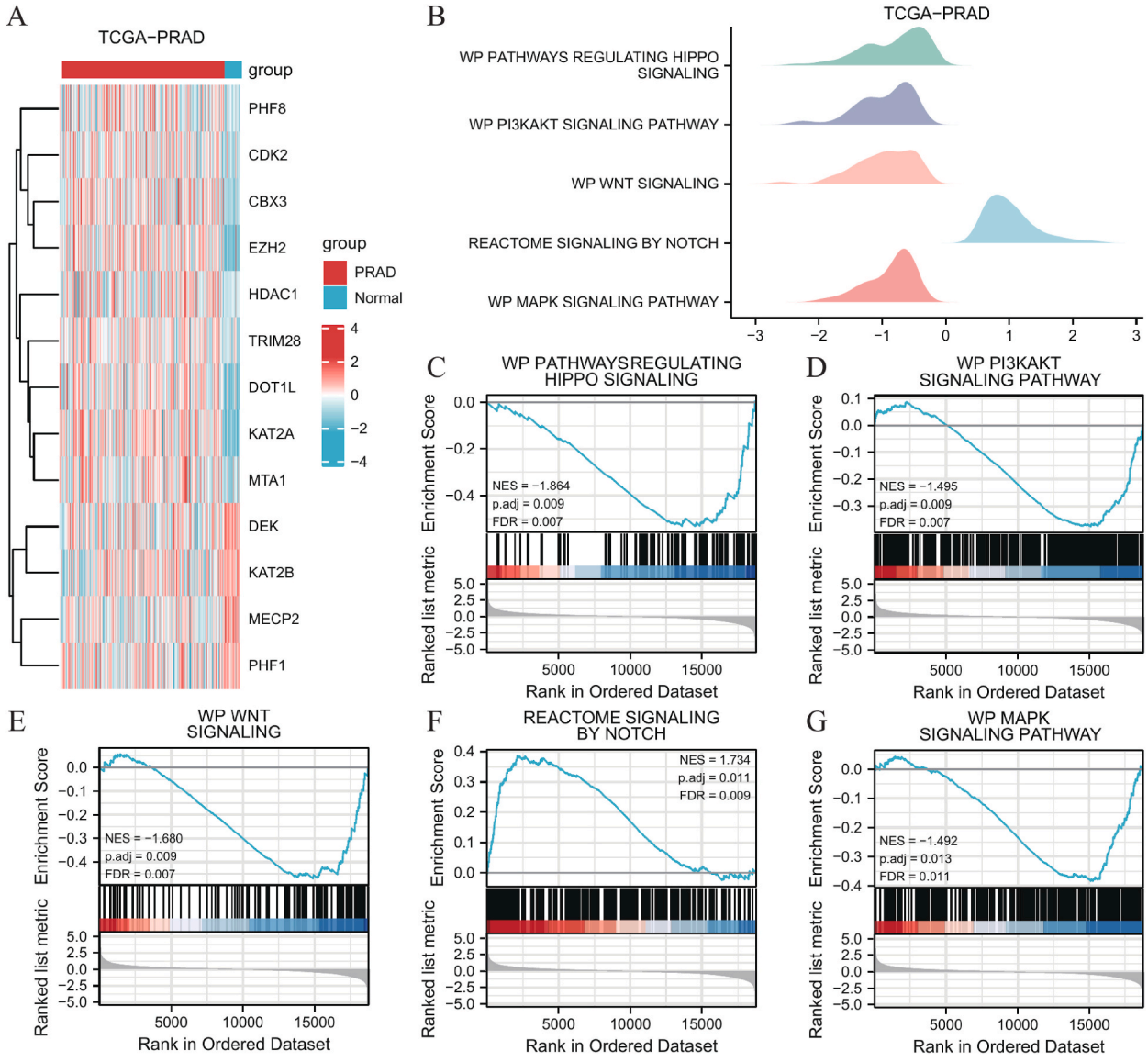


Fig. 3. Heatmap and GSEA analysis of ERGs.
A. Heatmap illustrating the expression levels of ERGs in the TCGA-PRAD dataset.
B. The top 5 major biological features enriched in the GSEA of the TCGA-PRAD dataset.
C-G. Significantly enriched pathways in the TCGA-PRAD dataset, including WP PATHWAYS REGULATING HIPPO SIGNALING, WP PI3KAKT SIGNALING PATHWAY, WP WNT SIGNALING, REACTOME SIGNALING BY NOTCH, and WP MAPK SIGNALING PATHWAY. The selection criteria are P.adj <0.05 and false discovery rate (FDR) value (q.value) < 0.25.

2.13. Statistical analysis

All data processing and analyses were performed using R software (version 4.20). Continuous variables are presented as mean ± standard deviation. The Wilcoxon rank-sum test was used to compare continuous variables between two groups, and an independent Student's t-test was used to assess the statistical significance of normally distributed variables. The Kruskal–Wallis test was used for comparisons involving three or more groups. Chi-square test or Fisher's exact test was used to compare and analyze the statistical significance of categorical variables between two groups. Cox regression analysis was used to evaluate the significance of the differences in survival time between the two patient groups. Unless otherwise specified, Spearman's correlation analysis was used to calculate the correlation coefficients between different molecules, and all P-values were two-tailed. Statistical significance was set at $P < 0.05$.

3. Results

3.1. Identification of differentially expressed genes (DEGs) and differentially expressed ERGs (DEERGs) in PRAD datasets

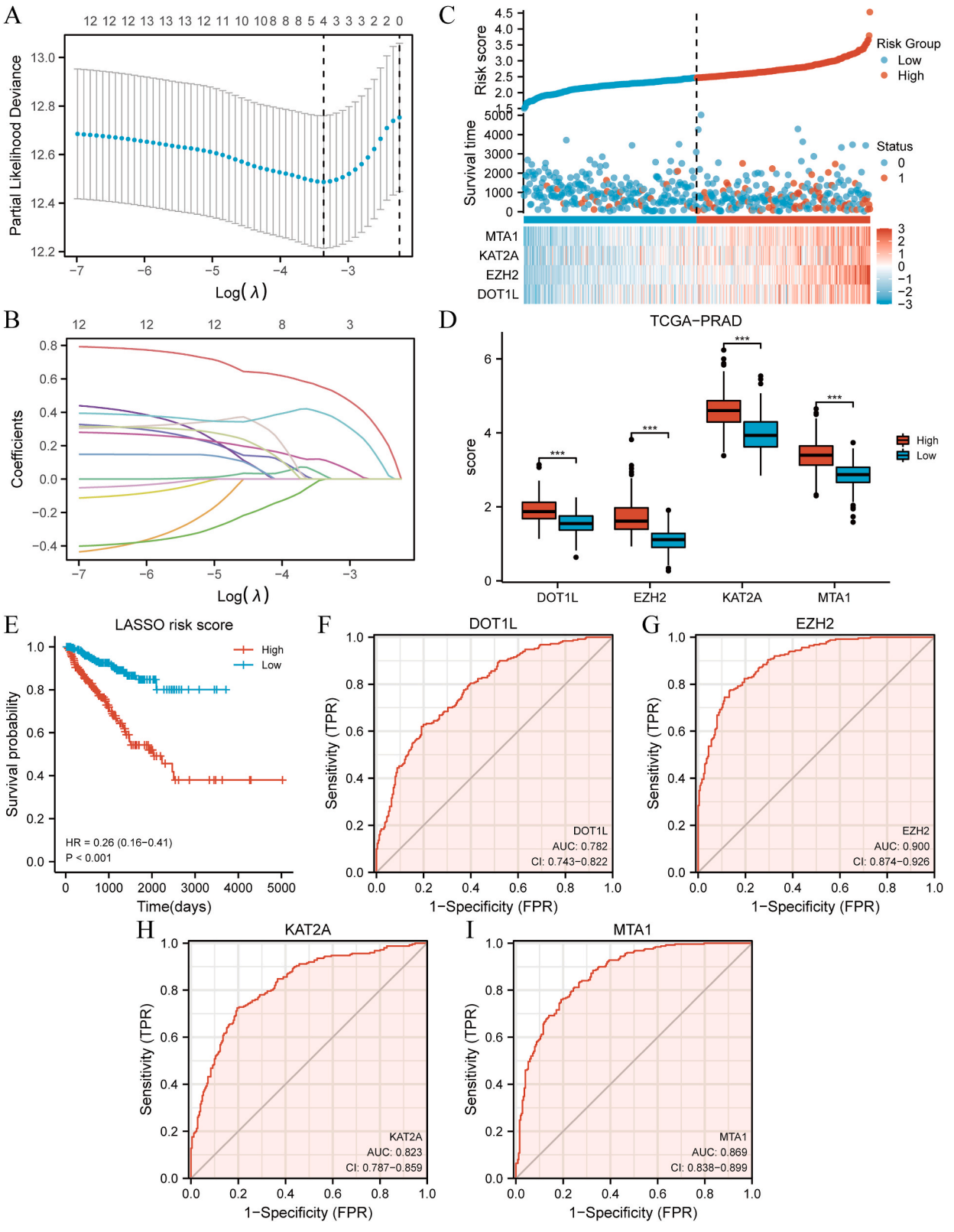
Fig. 1 shows a flowchart of this study. Batch effect correction and normalization were performed on the GEO datasets (GSE6919, GSE46602, GSE69223, and GSE141551), including 619 PRAD samples and 110 adjacent tissues. The effectiveness of batch-effect removal was assessed by comparing the pre- and post-batch-effect-removal datasets using box and PCA plots (Fig. 1A–D). The plots indicate that the batch effects in the PRAD dataset were largely eliminated. Furthermore, 501 PRAD samples and 52 paracancer samples (control) were obtained from TCGA database for transcriptomic data. The DEGs between the tumor and control groups were screened, which revealed 11656 DEGs, including 5992 upregulated and 5664 downregulated genes. Volcano plots were generated for the identified genes (Fig. 2A). Next, we drew a Venn diagram to identify the intersection with epigenetic related genes; 37 differentially expressed genes were obtained, including 22 upregulated genes (CBX3, CDK2, CHD4, DNMT3A, DNMT3B, DOT1L, EED, EHMT2, EZH2, HDAC1, KAT2A, KDM1A, MBD3, MTA1, MTA2, PHF8, SIRT6, SUV39H1, TDG, TET3, TRIM28, UHRF1) and 15 downregulated genes (BMI1, CTCF, DEK, HDAC5, KAT2B, MECP2, MPHOSPH8, PHF1, RB1, RBBP7, SIN3A, SIN3B, SIRT1, SMARCA5) (Fig. 2B and C). Subsequently, the expression trends and statistical significance of these 37 ERGs were analyzed using group comparison plots in both TCGA-PRAD and PRAD dataset (Fig. 2D and E). The results showed that 13 ERGs (CBX3, CDK2, DEK, DOT1L, EZH2, HDAC1, KAT2A, KAT2B, MECP2, MTA1, PHF1, PHF8, TRIM28) consistently exhibited significant differential expression and directionality in both datasets.

3.2. Correlation analysis and functional enrichment analysis of DEERGs

Based on the expression of the 13 ERGs in TCGA-PRAD dataset, correlation analysis was performed using the Spearman algorithm to examine the correlation between the expression levels of these 13 ERGs. The results are presented as a correlation heatmap (Fig. S2A) and a correlation chord diagram (Fig. S2B). From the correlation heatmap, we selected the four most representative gene pairs and visualized their correlation analysis results using correlation scatter plots (Figs. S2C–F). These correlation scatter plots revealed significant positive correlations between the expression levels of KAT2A and MTA1 ($r = 0.637$, $P < 0.001$, Fig. S2C), as well as CBX3 and EZH2 ($r = 0.586$, $P < 0.001$, Fig. S2D) in TCGA-PRAD dataset. However, significant negative correlations were observed between the expression levels of EZH2 and MECP2 ($r = -0.450$, $P < 0.001$, Fig. S2E) and between KAT2B and TRIM28 ($r = -0.432$, $P < 0.001$, Fig. S2F) in TCGA-PRAD dataset.

A heatmap was used to illustrate the expression levels of the 13 ERGs (CBX3, CDK2, DEK, DOT1L, EZH2, HDAC1, KAT2A, KAT2B, MECP2, MTA1, PHF1, PHF8, and TRIM28) in TCGA-PRAD dataset (Fig. 3A). Subsequently, GSEA was performed to investigate enriched pathways in this dataset. Significantly enriched gene sets were identified based on the criteria of P.adj <0.05 and false discovery rate (FDR) q-value <0.25. The results revealed significant enrichment of genes in pathways such as WP pathways regulating Hippo signaling (Fig. 3C), WP PI3KAKT signaling pathway (Fig. 3D), WP WNT signaling (Fig. 3E), reactome signaling by NOTCH (Fig. 3F), and the WP MAPK signaling pathway (Fig. 3G) (Supplemental Table 4). These results were further visualized using ridge plots (Fig. 3B) and pathway diagrams (Fig. 3C–G). GO functional enrichment analysis was conducted to analyze the biological processes, molecular functions, cellular components, and biological pathways associated with the 13 ERGs in relation to PRAD (Supplemental Table 5). The enrichment criteria for selecting the significant terms were set as described above. The results revealed that the 13 ERGs were mainly enriched in biological processes such as negative regulation of gene expression, epigenetics (GO:0045814), positive regulation of gene expression, epigenetics (GO:0045815), and covalent chromatin modification (GO:0016569). Cellular components were enriched in the nuclear chromatin (GO:0000790), chromosomal region (GO:0098687), and heterochromatin (GO:0000792). Additionally, their molecular functions were enriched in transcription corepressor activity (GO:0003714), histone deacetylase binding (GO:0042826), and transcription coactivator activity (GO:0003713) (Figs. S3A–B).

A combined GO and KEGG enrichment analysis was conducted by incorporating the logFC values of the 13 ERGs. We calculated the z-scores for each gene based on their logFC values and constructed a chord diagram (Fig. S3C) and a circular plot (Fig. S3D) to visualize the results. The circular plot highlights nuclear chromatin (GO:0000790) as a significantly upregulated cellular component, and the representative KEGG pathway, the Notch signaling pathway (hsa04330), is depicted (Fig. S3E).



(caption on next page)

Fig. 4. LASSO prognostic modeling of ERGs in TCGA-PRAD.
A. LASSO regression prognostic model plot for ERGs.
B–C. Risk factor plot and variable trajectory plot of the LASSO regression prognostic model.
D. Group comparison plot illustrating the expression levels of ERGs between the high-risk and low-risk groups.
E. Kaplan-Meier curves of PFI events between the high-risk and low-risk.
F–I. ROC curves of DOT1L (F), EZH2 (G), KAT2A (H), and MTA1 (I) between the high-risk and low-risk groups in the TCGA-PRAD dataset. *** indicate $P < 0.001$.

3.3. Construction the ERG-related risk model

To determine the prognostic value of the 13 ERGs in TCGA-PRAD dataset, we used LASSO regression analysis based on gene expression levels to construct a prognostic model (Fig. 4A). The model includes four ERGs (DOT1L, EZH2, KAT2A, and MTA1). The LASSO results were visualized using a LASSO variable trace plot (Fig. 4B). Next, we excluded normal samples from TCGA-PRAD dataset and divided the remaining cancer samples into high- and low-risk groups based on the median score of the samples in the model (Fig. 4C). Subsequently, we observed a highly significant statistical difference in the expression levels of these four ERGs between the high- and low-risk groups (Fig. 4D, $P < 0.001$). To assess the clinical relevance, survival curves were generated using PFI data to stratify patients into high- and low-risk groups (Fig. 4E). The findings revealed a significant difference ($P < 0.001$), indicating that patients in the high-risk category exhibited a notably accelerated rate of disease progression compared with those in the low-risk category. Furthermore, ROC curves were constructed for the four identified ERGs to discriminate between the high- and low-risk groups (Fig. 4F–I). The outcomes illustrated that EZH2 (AUC = 0.900; Fig. 4G) exhibited exceptional accuracy in predicting the outcome dichotomy between the high- and low-risk groups. Conversely, DOT1L (AUC = 0.782, Fig. 4F), KAT2A (AUC = 0.823, Fig. 6H), and MTA1 (AUC = 0.869, Fig. 4I) exhibited moderate accuracy in predicting outcome dichotomy between the two groups.

The expression levels and diagnostic efficacy of these four pivotal genes were validated using the GSE116918 dataset. The results indicated highly significant differences in the expression of these four ERGs between the high- and low-risk groups (Fig. S4A). Notably, EZH2 (AUC = 0.900, Fig. S4C) and MTA1 (AUC = 0.858, Fig. S4E) demonstrated high accuracy in predicting the outcomes in both groups. However, our diagnostic model exhibited suboptimal performance for prognostic prediction in the GSE116918 dataset (Figs. S4F–G), which was possibly attributable to the inherent heterogeneity observed in PRAD.

3.4. Identification of ERG clusters in PRAD

We analyzed the expression data of 13 ERGs in 499 samples from TCGA-PRAD dataset to construct consensus clustering. Based on their cumulative distribution function and δ area, we selected k = 2, where the ERGs appeared to be stably clustered (Fig. 5A and B). Subsequently, two epigenetic-related subtypes designated as cluster 1 and cluster 2 were obtained (Fig. 5C); cluster 2 exhibited a significant correlation with a superior prognosis, whereas cluster 1 showed the poorest survival probability (Fig. 5K). Next, we verified the differential expression of these 13 ERGs between the two PRAD subtypes in TCGA dataset (Fig. 5D). In addition to PHF1 and PHF8, the expression of the other 11 ERGs was significantly different ($P < 0.001$) between clusters 1 and 2. In addition, the PCA plot based on the two PRAD subtypes demonstrated a clear separation between them (Fig. 5E).

To explore the correlation between the expression of these 13 ERGs and different PRAD subtypes, we used ROC curves to predict the probability of each molecular subtype (Fig. 5F–J). The results indicated that KAT2A (AUC = 0.914) exhibited high accuracy in predicting the different PRAD subtypes. CBX3 (AUC = 0.784), CDK2 (AUC = 0.703), DOT1L (AUC = 0.833), EZH2 (AUC = 0.825), KAT2B (AUC = 0.778), MTA1 (AUC = 0.825), and TRIM28 (AUC = 0.758) demonstrated moderate accuracy in predicting different PRAD subtypes. However, DEK (AUC = 0.622), HDAC1 (AUC = 0.609), MECP2 (AUC = 0.609), PHF1 (AUC = 0.538), and PHF8 (AUC = 0.501) showed lower accuracy in predicting the different PRAD subtypes.

Furthermore, we investigated the differences in the hallmark gene sets between the two PRAD subtypes. GSVA was performed on the expression of all genes in the PRAD dataset. GSVA revealed statistically significant differential expression of 44 hallmark gene sets between clusters 1 and 2 ($P < 0.05$, Fig. S5A). To illustrate this further, we selected seven representative hallmark gene sets and plotted their expression patterns (Fig. S5B).

3.5. Construction of ERG-related scores and immune-related analysis

The ssGSEA algorithm was used to obtain ERG scores for each sample in TCGA-PRAD dataset based on the expression levels of the 13 ERGs. Subsequently, normal samples were excluded from the dataset and divided into high and low ERG score groups based on the median ERG score. We then constructed a group comparison plot to illustrate the differences in the expression of these 13 ERGs between the high and low ERG score groups in TCGA-PRAD dataset (Fig. 6A). The expression differences in 10 ERGs (CBX3, CDK2, DOT1L, EZH2, HDAC1, KAT2A, KAT2B, MTA1, PHF8, and TRIM28) were statistically significant ($P < 0.05$). Survival analysis based on PFI data for the high and low ERG score groups showed that the difference was highly significant ($P = 0.009$). Patients in the high ERG score group experienced faster disease progression than those in the low ERG score group (Fig. 6B). The correlation between the 13 ERGs and ERG score groups was analyzed using ROC curves (Fig. 6C–G). Except for poor prediction performance exhibited by DEK (AUC = 0.550), HDAC1 (AUC = 0.591), KAT2B (AUC = 0.561), MECP2 (AUC = 0.491), PHF1 (AUC = 0.492), and TRIM28 (AUC = 0.642), the remaining ERGs had a certain accuracy (CBX3 (AUC = 0.727), CDK2 (AUC = 0.776), DOT1L (AUC = 0.753), EZH2 (AUC = 0.790), KAT2A (AUC = 0.736), MTA1 (AUC = 0.756), and PHF8 (AUC = 0.730)) in predicting the ERG score groups.

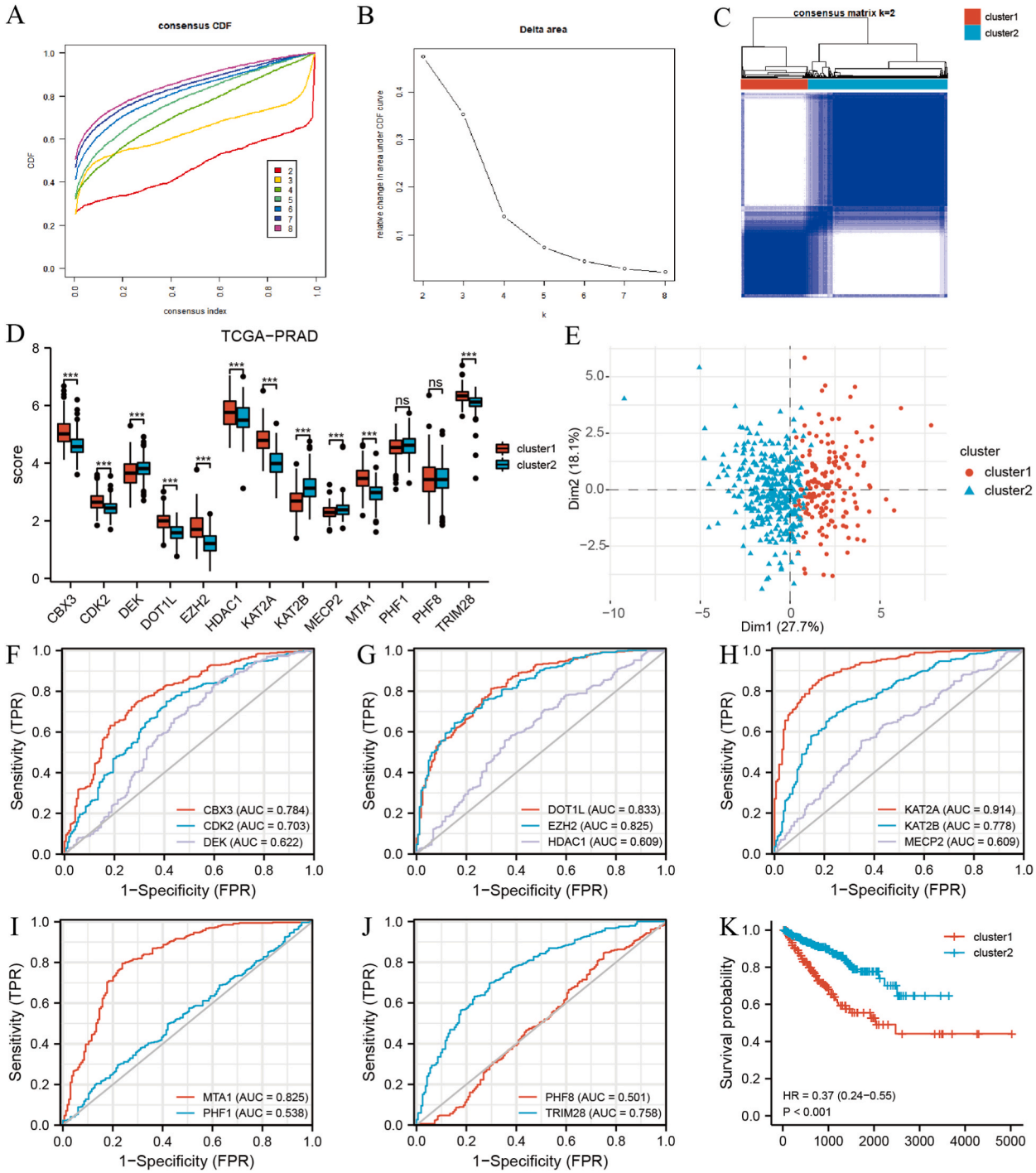


Fig. 5. Consensus clustering establishes disease subtypes for PRAD.
A-B. Cumulative distribution function (CDF) plots (A) and Delta area under the CDF curve plots (B) for different numbers of PRAD subtypes.
C. Consensus clustering (K = 2) results plot for PRAD.
D. PRAD subtypes cluster1 and cluster2 exhibit differential expression of ERGs.
E. Principal Component Analysis of PRAD subtypes cluster1 and cluster2.
F-J. ROC curve results for ERGs in the TCGA-PRAD dataset based on disease subtype grouping (cluster1 and cluster2).
K. Kaplan-Meier curves of PFI events between disease subgroups (cluster1 and cluster2) in the TCGA-PRAD dataset. *** indicate $P < 0.001$.

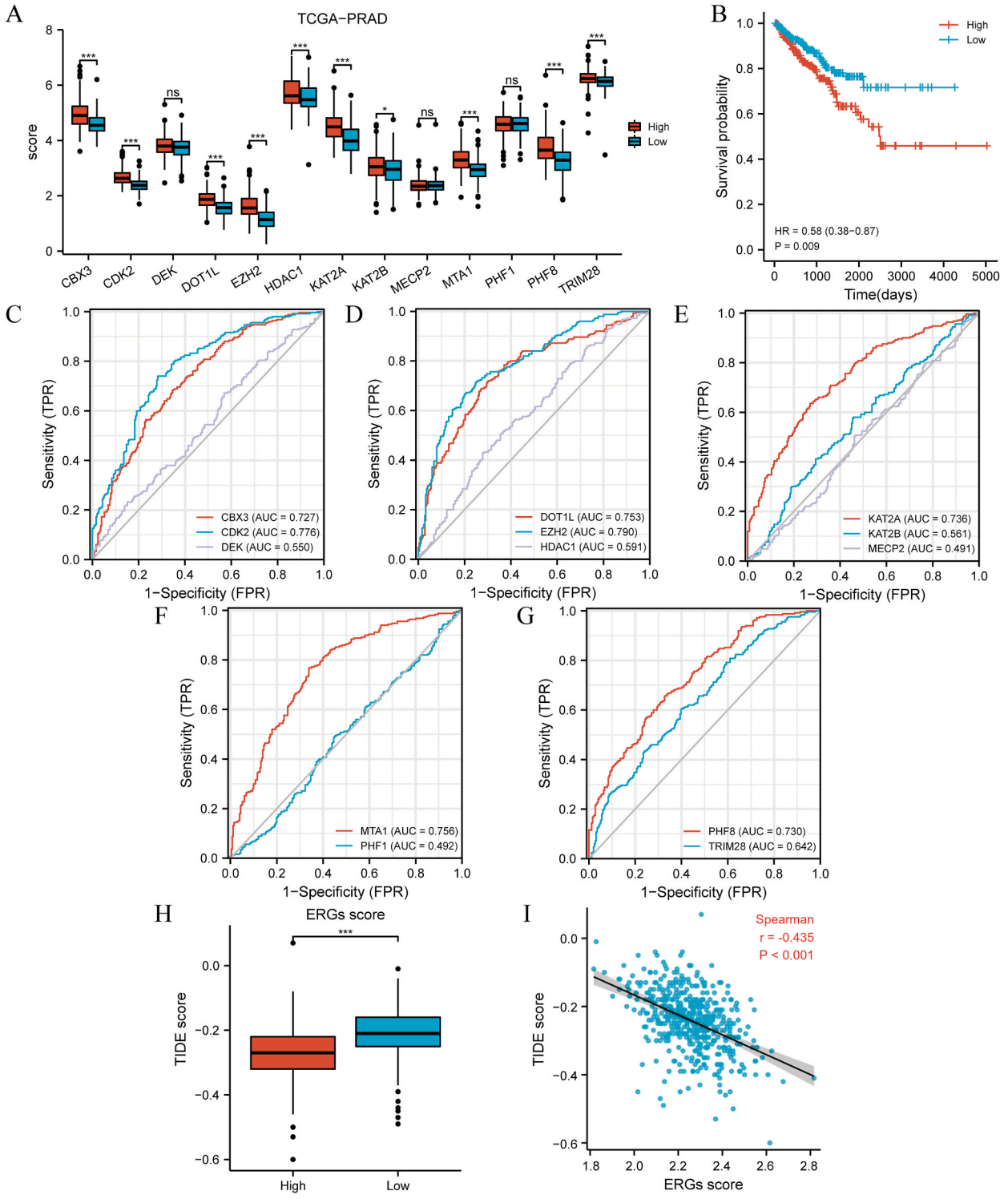


Fig. 6. Construction of ERGs scores and immunotherapy analysis.
A. The expression levels of ERGs between the high and low ERGs score groups in the TCGA-PRAD dataset.
B. Kaplan-Meier curves of PFI events between the high and low ERGs score groups in the TCGA-PRAD dataset.
C-G. ROC curve results for ERGs in the TCGA-PRAD dataset based on high and low ERGs score groupings.
H. Group comparison plot showing TIDE scores between the high and low ERGs score groups in the TCGA-PRAD dataset.
I. Scatter plot illustrating the correlation between ERGs scores and TIDE scores in the TCGA-PRAD dataset. *** indicate $P < 0.001$.

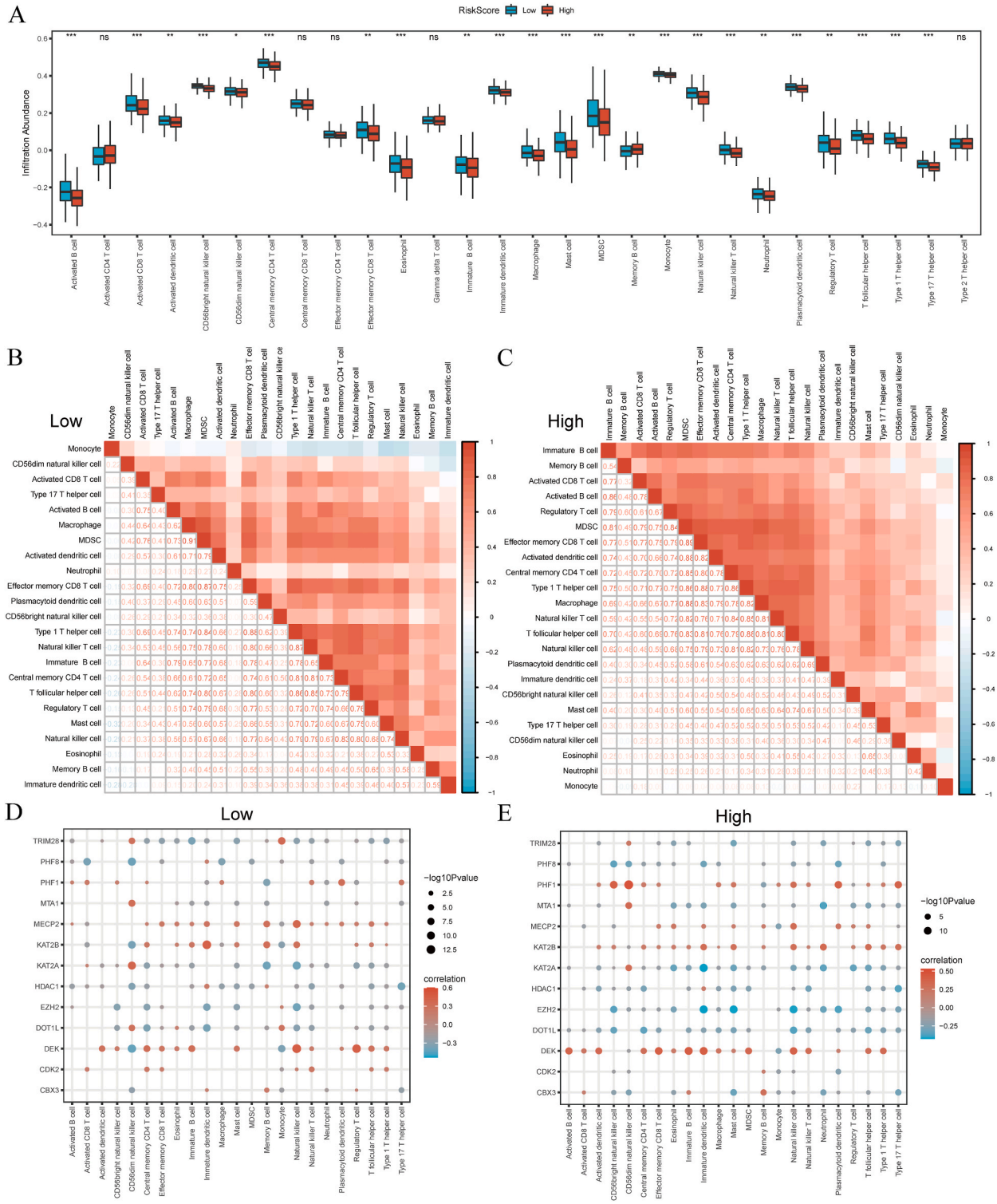


Fig. 7. Immune infiltration between groups with high and low ERGs scores on the TCGA-PRAD dataset.

A. The results of ssGSEA immune infiltration analysis between the high and low ERGs score groups in the TCGA-PRAD dataset.

B–C. Correlation analysis of immune cell infiltration abundance in the low ERGs score group (B) and high ERGs score group (C) of the TCGA-PRAD dataset.

D-E. Correlation heatmap between immune cells and ERGs in the low ERGs score group (D) and high ERGs score group (E) of the TCGA-PRAD dataset. * indicate $P < 0.05$, ** indicate $P < 0.01$, *** indicate $P < 0.001$.

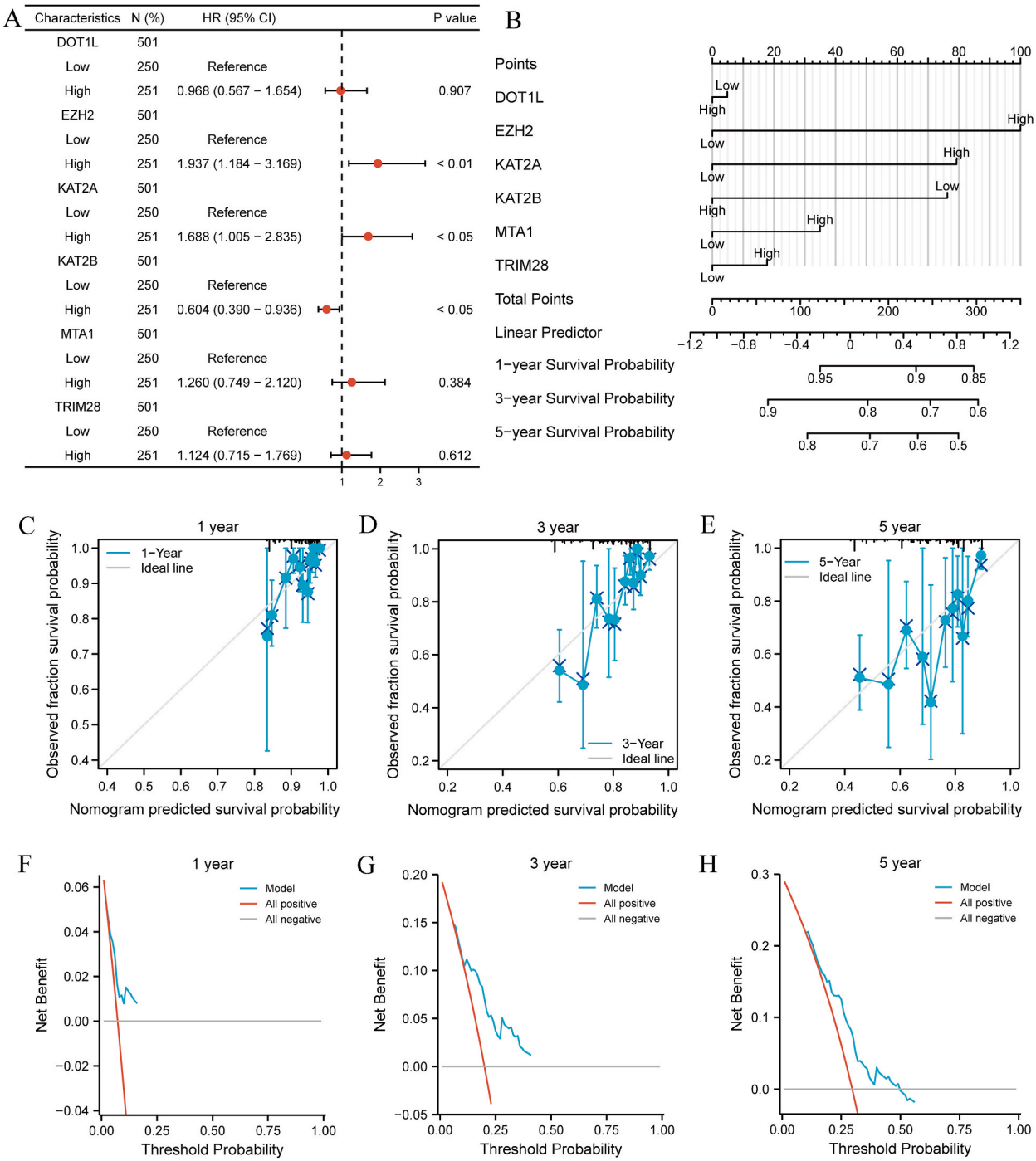


Fig. 8. Clinical Prognosis Analysis.

A-B. Forest plot (A) and column chart (B) depicting the results of multivariable Cox regression analysis.

C-E. Calibration curve plots for 1-year (C), 3-year (D), and 5-year (E) time frames based on the nomogram analysis of the multivariable Cox regression model.

F-G. DCA plots for the 1-year (F), 3-year (G), and 5-year (H) Cox regression prognostic models. The blue line representing the 1-year prognosis is closest to the ideal gray line, indicating a better predictive performance of the model for the first year compared to the 3-year and 5-year prognoses. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

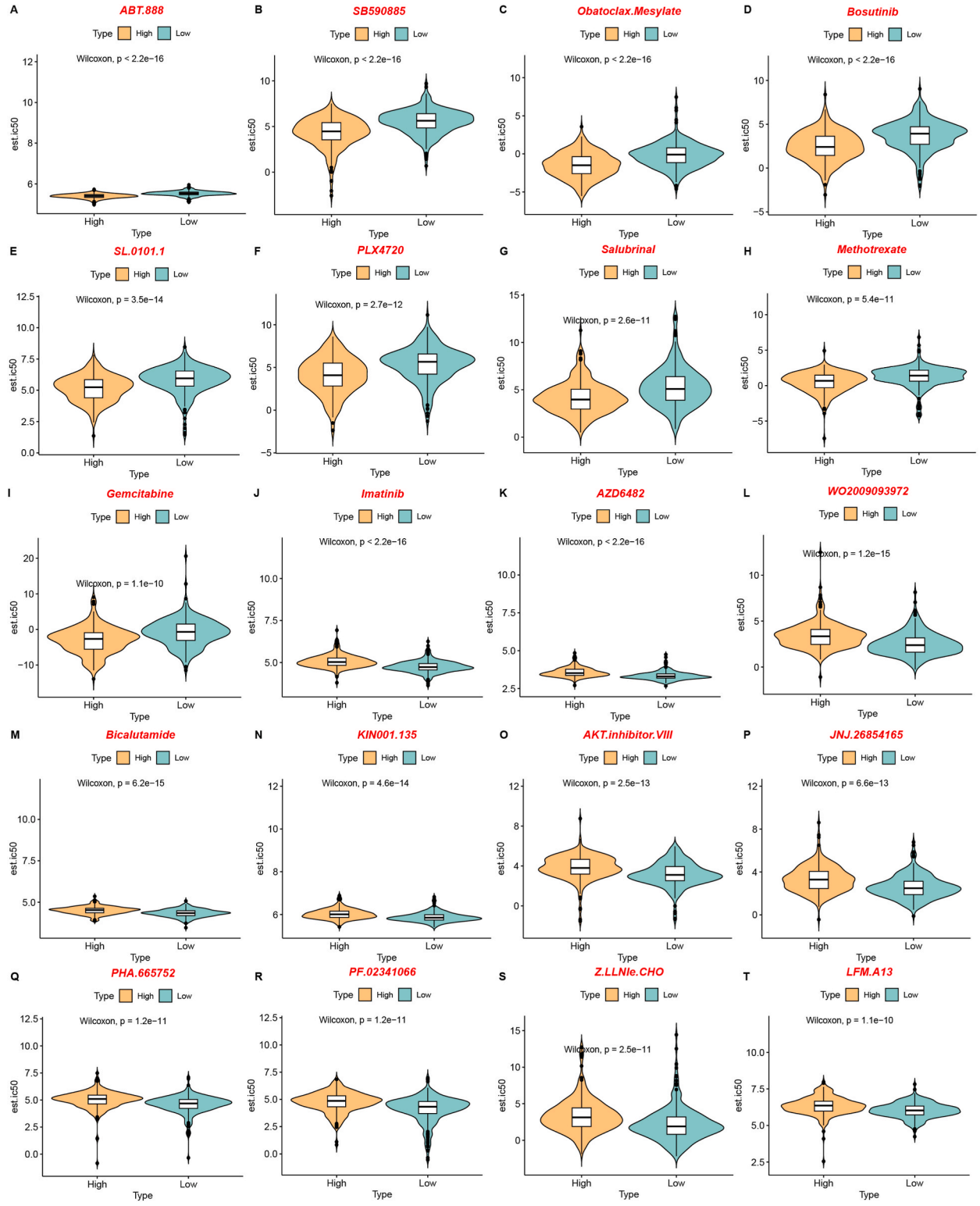


Fig. 9. Value of the risk score for predicting drug sensitivity.
Drug sensitivity in the high- and low-risk groups was estimated using the oncoPredict algorithm in the TCGA-PRAD dataset. ABT.888 (A), SB590885 (B), Obatoclax.Mesylate (C), Bosutinib (D), SL.0101.1 (E), PLX4720 (F), Salubrinal (G), Methotrexate (H), Gemcitabine (I), Imatinib (J), AZD6482 (K), WO2009093972 (L), Bicalutamide (M), KIN001.135 (N), AKT.inhibitor.VIII (O), JNJ.26854165 (P), PHA.665752 (Q), PF.02341066 (R), Z. LLNle.CHO (S), LFM.A13 (T).

Considering the significant role of immunotherapy in PRAD, we evaluated the TIDE scores for each patient in TCGA dataset, which represents sensitivity to immunotherapy. The results showed a difference in TIDE scores between the ERG score groups in patients with PRAD ($P < 0.01$), with the low ERG score group having higher TIDE scores, indicating a poorer immunotherapy response than the high ERG score group (Fig. 6H). Additionally, correlation analysis showed a weak negative correlation between TIDE and ERG scores in patients with PRAD (Fig. 6I, $r = -0.435$, $P < 0.001$).

To investigate the differences in immune cell infiltration between the high and low ERG score groups in TCGA-PRAD dataset, we performed ssGSEA to calculate the infiltration abundance of 28 immune cell types. The results demonstrated significant differences ($P < 0.05$) in the infiltration abundance of 23 immune cell types between the high and low ERG score groups in TCGA-PRAD dataset

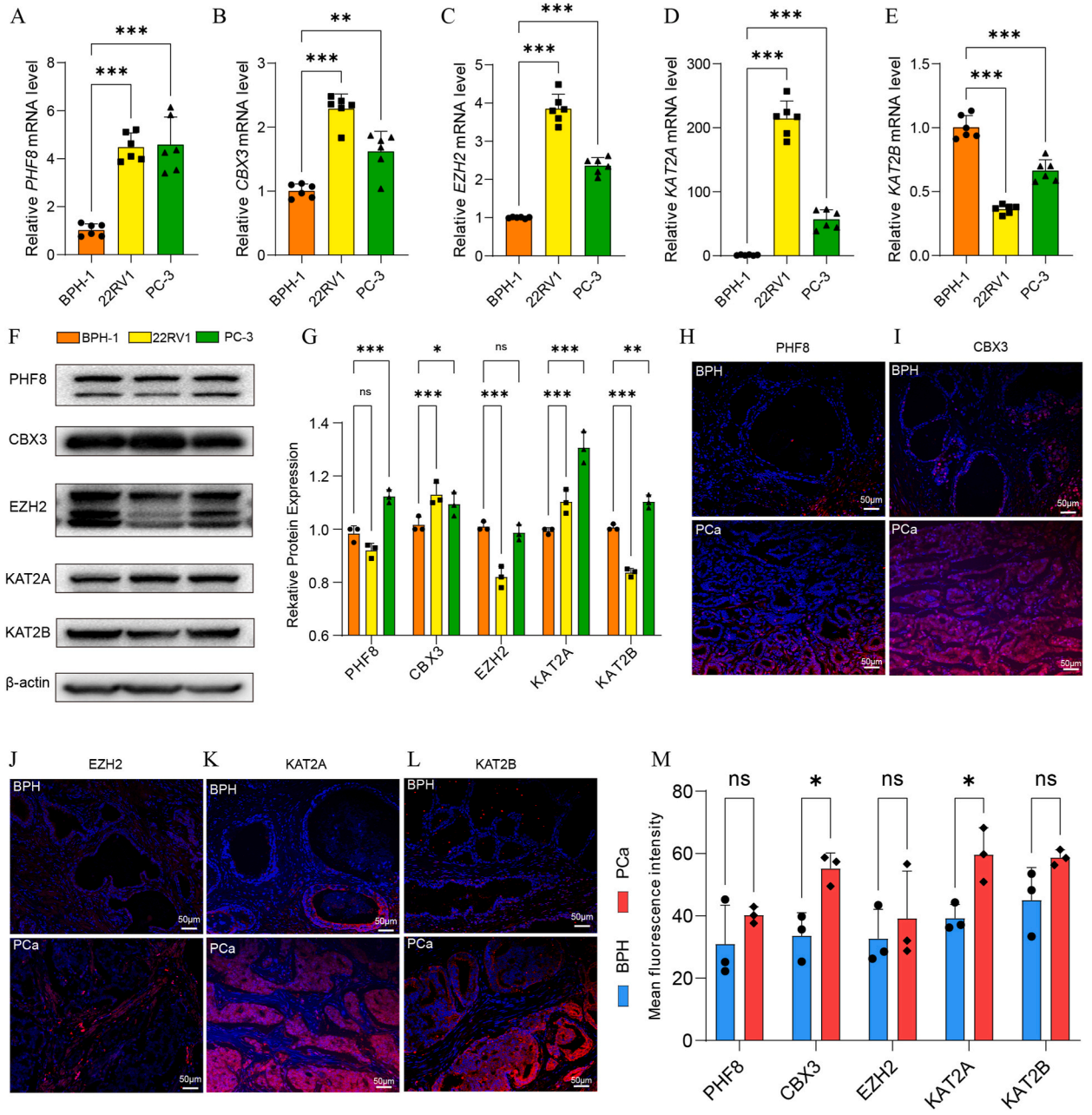


Fig. 10. Experimental Validation.

A-E. RT-qPCR analysis of ERGs mRNA expression in BPH-1 cells, 22RV1 cells, and PC-3 cells. Data represent Mean ± SD, n = 6.

F-G. Western blot and statistics analysis of ERGs protein expression in BPH-1 cells, 22RV1 cells, and PC-3 cells. Data represent Mean ± SD, n = 3.

H-M. immunofluorescence detection and statistics analysis of the expression of CBX3 and KAT2A in clinical samples. Data represent Mean ± SD, n = 3. ** indicate $P < 0.01$, *** indicate $P < 0.001$.

(Fig. 7A). These immune cell types included activated B cells, activated CD8 T cells, activated dendritic cells, CD56bright natural killer cells, CD56dim natural killer cells, central memory CD4 T cells, effector memory CD8 T cells, eosinophils, immature B cells, immature dendritic cells, macrophages, mast cells, myeloid-derived suppressor cells, memory B cells, monocytes, natural killer cells, natural killer T cells, neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells, type 1 T helper cells, and type 17 T helper cells. Furthermore, correlation analysis revealed correlations among the 25 immune cell types in TCGA-PRAD dataset (Fig. 7B and C). We then investigated the correlations between the infiltration abundance of these 23 immune cell types and the expression levels of 13 ERGs within the low (Fig. 7D) and high ERG score groups (Fig. 7E). In the low ERG score group, we observed significant positive correlations between DEK and MECP2 expression in these 23 immune cell types. In the high ERG score group, significant positive correlations were observed between DEK and KAT2B and these 23 immune cell types, whereas DOT1L, EZH2, and KAT2A showed significant negative correlations with these immune cell types.

3.6. Construction of nomogram and calibration curves

Univariate and multivariate Cox regression analyses were performed to analyze the expression levels of the 13 ERGs in TCGA-PRAD dataset and explore their associations with clinical prognosis. First, we conducted univariate Cox regression analysis on the expression levels of the 13 ERGs, and factors with $P < 0.05$ were selected for inclusion in the multivariate Cox regression analysis. A multivariate Cox regression model was constructed (Supplemental Table 5) and a forest plot was generated (Fig. 8A) to present the results (Supplemental Table 6). Additionally, a nomogram for predicting the 1-, 3-, and 5-year overall survival (OS) of patients with PRAD was developed based on these six significant genes (Fig. 8B). The total score was calculated for each patient by summing the points assigned to each variable. Subsequently, the 1-, 3-, and 5-year OS probabilities were estimated by projecting the total score onto the corresponding nomogram scale. Furthermore, we conducted a calibration analysis of the nomogram of the multivariate Cox regression model for the 1-year (Fig. 8C), 3-year (Fig. 8D), and 5-year (Fig. 8E) prognoses, and calibration curve plots were generated. DCA was used to evaluate the clinical utility of the constructed Cox regression prognostic model for the 1-year (Fig. 8F), 3-year (Fig. 8G), and 5-year (Fig. 8H) timeframes. The blue line representing the model's stability indicates better predictive performance of the model for the 3-year than for the 1- and 5-year outcomes.

3.7. Risk model prediction of drug sensitivity

To assess the predictive capability of our risk score model for drug sensitivity, we conducted a comparative analysis of drug sensitivity between the high and low ERG risk groups using the OncoPredict algorithm. As depicted in Fig. 9, the high-risk group exhibited notably increased sensitivity to ABT-888, SB590885, obatoclax mesylate, bosutinib, SL01011, PLX4720, salbrinal, methotrexate, and gemcitabine (Fig. 9A–I). Conversely, the low-risk group demonstrated an increased sensitivity to imatinib, AZD6482, WO2009093972, bicalutamide, KIN001.135, AKT inhibitor VIII, JNJ.26854165, PHA.665752, PF.02341066, Z.LLNle.CHO, and LFM.A13 (Fig. 9J-T).

3.8. Experiment validation

RT-qPCR was performed to validate ERG expression in BPH-1, 22RV1, and PC-3 cells. Our results demonstrated the upregulation of PHF8, CBX3, EZH2, and KAT2A expression in tumor cells, whereas KAT2B expression was downregulated (Fig. 10A–E). Given the extensive research conducted on these four genes in tumors, it is imperative to further investigate their expression and function in PRAD (Supplemental Table S7). Furthermore, immunoblotting confirmed the increased expression of CBX3 and KAT2A in PCa cells (Fig. 10F). Immunofluorescence staining of clinical PCa samples revealed a significant increase in CBX3 and KAT2A expression, suggesting their potential biological significance (Fig. 10G–K). These findings highlight the need for further investigations into the functional roles of CBX3 and KAT2A in PCa.

4. Discussion

Recent advances in epigenomic research have provided new insights into the pathogenesis and potential therapeutic targets in a wide range of pathological conditions, including PCa [23,24]. Kathleen et al. identified 1178 differentially methylated loci in intermediate-risk PCa relative to benign tissue [25]. This seminal finding underscores the crucial role of epigenetic modification of DNA methylation in metastatic PCa development [26]. In this study, we identified 13 ERGs associated with PRAD prognosis through a comprehensive analysis of TCGA and GEO datasets. Based on these 13 key genes, we conducted epigenetic scoring of PRAD samples from TCGA, which revealed that patients with higher scores exhibited enhanced responses to immunotherapy. This study may assist in our understanding of the epigenetic status and antitumor immune response in PRAD and thus provide potential targets for clinical therapeutic interventions.

CBX3 plays a crucial role in the epigenetic regulation of cancer development. Notably, CBX3 expression is elevated in PRAD and has shown superiority over the Gleason score as a predictor of biochemical recurrence after radical prostatectomy [27]. Our bioinformatics analysis indicated that increased CBX3 expression is closely associated with PCa progression and predicts an unfavorable prognosis. Concurrently, our experimental results corroborated the substantial elevation in both the mRNA and protein levels of CBX3 in PCa cell lines. Furthermore, this expression trend was consistently reproduced during the validation using clinical specimens. Therefore, CBX3 has significant potential as a therapeutic target for the epigenetic regulation of PRAD.

KAT2A and KAT2B are histone acetyltransferases that primarily act as transcriptional activators. However, the association between KAT2B and PRAD has not yet been reported. Lu et al. reported that KAT2A facilitates the translocation of androgen receptors into the nucleus, thereby promoting abiraterone resistance in castration-resistant PCa [28]. Our data revealed a significant increase in KAT2A expression in both PCa cells and clinical samples. However, the mechanisms underlying resistance to PRAD in this context warrant further investigation.

In recent years, numerous somatic and germline alterations have been identified in the Hippo, PI3K, and MAPK signaling pathways, indicating potential factors that contribute to disease initiation, metastasis, and castration resistance [29–31]. Additionally, elevated expression of the epigenetic regulator HMGA2 within the tumor microenvironment is associated with the induction of prostatic intraepithelial neoplastic lesions, presumably by enhancing paracrine Wnt signaling in adjacent prostatic epithelial cells [32]. Our current investigation revealed a noteworthy attenuation of Hippo signaling activity within the framework of PRAD. The Hippo-YAP/TAZ signaling axis is an attractive candidate for prospective therapeutic interventions in PCa [33]. Analogous to the Hippo signaling cascade, the NOTCH signaling pathway also maintains evolutionary conservation [34]. The Notch signaling pathway plays a pivotal role in the development, progression, and metastasis of PCa [35]. This study clearly illustrated a notable elevation in the Notch pathway in the context of PRAD. Collectively, these findings suggest that interventions targeting the NOTCH pathway could hold promise as an adjuvant therapeutic approach.

Another significant discovery underscores the prospect of repurposing established pharmaceuticals to enhance the treatment efficacy in patients with PRAD. Our analysis predicted that certain drugs could exhibit enhanced therapeutic efficacy in patients with high risk scores, whereas others demonstrated improved effectiveness in those with low risk scores. Importantly, some of these drugs have previously been validated for their therapeutic potential against PRAD. For instance, ABT-888 (veliparib), imatinib, gemcitabine, and obatoclax mesylate have shown promising results in inhibiting PCa progression. Additionally, bicalutamide has gained approval as an orally active nonsteroidal androgen receptor antagonist and is used as a first-line treatment, particularly in combination with luteinizing hormone-releasing hormone analogs or surgical orchiectomy, for the management of advanced PCa.

Our findings hold great potential for guiding future PCa research by identifying new therapeutic targets and prognostic biomarkers and ultimately advancing personalized medical approaches. By understanding the genetic and molecular differences among patients, treatment strategies can be tailored to improve patient outcomes and quality of life. Specifically, the modulation and targeting of epigenetic-related genes, such as those examined in this study, offer promise for mitigating PCa progression by influencing crucial cellular processes and signaling pathways, including the Hippo, PI3K, and Wnt pathways. Additionally, the intricate interplay between these genes and tumor-associated immune cells underscores their importance in PCa management. This comprehensive exploration of epigenetics in PRAD, including the analysis of key genes, pathways, and immune cell infiltration, provides a foundation for novel therapeutic approaches with multifaceted effects encompassing molecular, signaling, and immune-related mechanisms.

5. Conclusions

This study provides crucial insights into the epigenetic landscape in PRAD, identifying 13 key genes associated with prognosis. Clustering and pathway analyses enhance our understanding of disease heterogeneity. The developed epigenetic scoring approach is a promising prognostic tool, with CBX3 and KAT2A validated as key players. These findings will pave the way for targeted and personalized interventions for PCa treatment.

Funding statement

Thanks for the financial support provided by the National Natural Science Foundation (No. 82303120), the Natural Science Foundation of Hunan Province (No. 2022JJ40747, No. 2022JJ30908), Innovation Guidance Project of Clinical Medical Technology of Hunan Province (No.2021SK53710), and the Wisdom Accumulation and Talent Cultivation Project of the Third Xiangya Hopsital of Central South University, China (No.20210312) for this research.

Data availability statement

No new datasets or code were generated in the current study. The datasets analyzed during this research are available in public repositories. The gene expression data can be accessed in TCGA and GEO databases under the following accession numbers: GSE46602, GSE69223, GSE6919, and GSE141551.

Ethical Statement

Prostate cancer specimens were obtained from patients who underwent radical prostatectomy at the Department of Urology, the Third Xiangya Hospital of Central South University, with adjacent paracancerous tissues serving as controls. This study adheres to the principles of the Declaration of Helsinki and was conducted in accordance with the ethical guidelines approved by the Third Xiangya Hospital's ethics committee (Approval No. 2023-S140).

CRediT authorship contribution statement

Youyou Li: Writing – original draft. **Chao Li:** Writing – review & editing, Data curation. **Longxiang Wu:** Visualization,

Methodology. **Jiaren Li:** Investigation, Data curation. **Yu Gan:** Visualization, Methodology. **Shuo Tan:** Investigation, Data curation. **Lei Zhou:** Investigation, Data curation. **Wei Xiong:** Investigation, Data curation. **Liang Zhou:** Conceptualization. **Cheng Li:** Conceptualization. **Jiahao Liu:** Conceptualization. **Dingwen Liu:** Resources, Investigation. **Yichuan Wang:** Resources, Investigation. **Yunlong Fu:** Resources, Investigation. **Kun Yao:** Writing – review & editing, Funding acquisition. **Long Wang:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

No.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e30941.

References

[1] O. Bergengren, K.R. Pekala, K. Matsoukas, J. Fainberg, S.F. Mungovan, O. Bratt, F. Bray, O. Brawley, A.N. Luckenbaugh, L. Mucci, et al., 2022 update on prostate cancer epidemiology and risk factors-A systematic review, Eur. Urol. 84 (2) (2023) 191–206. https://10.1016/j.eururo.2023.04.021.
[2] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249. https://10.3322/caac.21660.
[3] C. Parker, E. Castro, K. Fizazi, A. Heidenreich, P. Ost, G. Procopio, B. Tombal, S. Gillessen, Prostate cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, Ann. Oncol. : official journal of the European Society for Medical Oncology 31 (9) (2020) 1119–1134. https://10.1016/j.annonc.2020.06.011.
[4] A.D. Valdez-Vargas, H.M. Sánchez-López, M.A. Badillo-Santoyo, R.E. Maldonado-Valadez, B.O. Manzo-Pérez, V.M. Pérez-Abarca, G. Manzo-Pérez, M. A. Vanzzini-Guerrero, J.A. Álvarez-Canales, Recurrence rate of localized prostate cancer after radical prostatectomy according to D'amico risk classification in a tertiary referral hospital: association study, Cirugía Cir. 89 (4) (2021) 520–527. https://10.24875/ciru.200007601.
[5] A. Farsetti, B. Illi, C. Gaetano, How epigenetics impacts on human diseases, Eur. J. Intern. Med. 114 (2023) 15–22. https://10.1016/j.ejim.2023.05.036.
[6] D. Hanahan, Hallmarks of cancer: new dimensions, Cancer Discov. 12 (1) (2022) 31–46. https://10.1158/2159-8290.CD-21-1059.
[7] J. López, A.M. Añazco-Guenkova, Ó. Monteagudo-García, S. Blanco, Epigenetic and epitranscriptomic control in prostate cancer, Genes 13 (2) (2022). https://10.3390/genes13020378.
[8] R. Ge, Z. Wang, R. Montironi, Z. Jiang, M. Cheng, M. Santoni, K. Huang, F. Massari, X. Lu, A. Cimadamore, et al., Epigenetic modulations and lineage plasticity in advanced prostate cancer, Ann. Oncol. : official journal of the European Society for Medical Oncology 31 (4) (2020) 470–479. https://10.1016/j.annonc.2020.02.002.
[9] S. Goel, V. Bhatia, T. Biswas, B. Ateeq, Epigenetic reprogramming during prostate cancer progression: a perspective from development, Semin. Cancer Biol. 83 (2022) 136–151. https://10.1016/j.semcancer.2021.01.009.
[10] A. Colaprico, T.C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, T.S. Sabedot, T.M. Malta, S.M. Pagnotta, I. Castiglioni, et al., TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, Nucleic Acids Res. 44 (8) (2016) e71. https://10.1093/nar/gkv1507.
[11] M.M. Mortensen, S. Hoyer, A.S. Lynnerup, T.F. Orntoft, K.D. Sorensen, M. Borre, L. Dyrskjot, Expression profiling of prostate cancer tissue delineates genes associated with recurrence after prostatectomy, Sci. Rep. 5 (2015) 16018. https://10.1038/srep16018.
[12] S. Meller, H.A. Meyer, B. Bethan, D. Dietrich, S.G. Maldonado, M. Lein, M. Montani, R. Reszka, P. Schatz, E. Peter, et al., Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer, Oncotarget 7 (2) (2016) 1421–1438. https://10.18632/oncotarget.6370.
[13] Y.P. Yu, D. Landsittel, L. Jing, J. Nelson, B. Ren, L. Liu, C. McDonald, R. Thomas, R. Dhir, S. Finkelstein, et al., Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy, J. Clin. Oncol. : official journal of the American Society of Clinical Oncology 22 (14) (2004) 2790–2799. https://10.1200/jco.2004.05.158.
[14] R. Rubicz, S. Zhao, C. April, J.L. Wright, S. Kolb, I. Coleman, D.W. Lin, P.S. Nelson, E.A. Ostrander, Z. Feng, et al., Expression of cell cycle-regulated genes and prostate cancer prognosis in a population-based cohort, Prostate 75 (13) (2015) 1354–1362. https://10.1002/pros.23016.
[15] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell systems 1 (6) (2015) 417–425. https://10.1016/j.cels.2015.12.004.
[16] Z.H. Wu, D.L. Yang, L. Wang, J. Liu, Epigenetic and immune-cell infiltration changes in the tumor microenvironment in hepatocellular carcinoma, Front. Immunol. 12 (2021) 793343. https://10.3389/fimmu.2021.793343.
[17] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (5) (2012) 284–287. https://10.1089/omi.2011.0118.
[18] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7. https://10.1186/1471-2105-14-7.
[19] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin. Epigenet. 11 (1) (2019) 123. https://10.1186/s13148-019-0730-1.
[20] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573. https://10.1093/bioinformatics/btq170.
[21] J. Fu, K. Li, W. Zhang, C. Wan, J. Zhang, P. Jiang, X.S. Liu, Large-scale public data reuse to model immunotherapy response and resistance, Genome Med. 12 (1) (2020) 21. https://10.1186/s13073-020-0721-z.
[22] L. Zhou, Y. Li, J. Li, H. Yao, J. Huang, C. Li, L. Wang, Decoding ceRNA regulatory network and autophagy-related genes in benign prostatic hyperplasia, Int. J. Biol. Macromol. 225 (2023) 997–1009. https://10.1016/j.ijbiomac.2022.11.162.
[23] K. Kukkonen, S. Taavitsainen, L. Huhtala, J. Uusi-Makela, K.J. Granberg, M. Nykter, A. Urbanucci, Chromatin and epigenetic dysregulation of prostate cancer development, progression, and therapeutic response, Cancers 13 (13) (2021). https://10.3390/cancers13133325.

[24] A. Kumaraswamy, K.R. Welker Leng, T.C. Westbrook, J.A. Yates, S.G. Zhao, C.P. Evans, F.Y. Feng, T.M. Morgan, J.J. Alumkal, Recent advances in epigenetic biomarkers and epigenetic targeting in prostate cancer, Eur. Urol. 80 (1) (2021) 71–81. https://10.1016/j.eururo.2021.03.005.
[25] K.E. Houlahan, Y.J. Shiah, A. Gusev, J. Yuan, M. Ahmed, A. Shetty, S.G. Ramanand, C.Q. Yao, C. Bell, E. O'Connor, et al., Genome-wide germline correlates of the epigenetic landscape of prostate cancer, Nat. Med. 25 (10) (2019) 1615–1626. https://10.1038/s41591-019-0579-z.
[26] S. Zhao, A. Leonardson, M.S. Geybels, A.S. McDaniel, M. Yu, S. Kolb, H. Zong, K. Carter, J. Siddiqui, A. Cheng, et al., A five-CpG DNA methylation score to predict metastatic-lethal outcomes in men treated with radical prostatectomy for localized prostate cancer, Prostate 78 (14) (2018) 1084–1091. https://10.1002/pros.23667.
[27] J. Slezak, M. Truong, W. Huang, D. Jarrard, HP1γ expression is elevated in prostate cancer and is superior to Gleason score as a predictor of biochemical recurrence after radical prostatectomy, BMC Cancer 13 (2013) 148. https://10.1186/1471-2407-13-148.
[28] D. Lu, Y. Song, Y. Yu, D. Wang, B. Liu, L. Chen, X. Li, Y. Li, L. Cheng, F. Lv, et al., KAT2A-mediated AR translocation into nucleus promotes abiraterone-resistance in castration-resistant prostate cancer, Cell Death Dis. 12 (8) (2021) 787. https://10.1038/s41419-021-04077-w.
[29] F. Koinis, E. Chantzara, M. Samarinas, A. Xagara, Z. Kratiras, V. Leontopoulou, A. Kotsakis, Emerging role of YAP and the Hippo pathway in prostate cancer, Biomedicines 10 (11) (2022). https://10.3390/biomedicines10112834.
[30] W. Abida, J. Armenia, A. Gopalan, R. Brennan, M. Walsh, D. Barron, D. Danila, D. Rathkopf, M. Morris, S. Slovin, et al., Prospective genomic profiling of prostate cancer across disease states reveals germline and somatic alterations that may affect clinical decision making, JCO Precis Oncol 2017 (2017). https://10.1200/po.17.00029.
[31] C.C. Pritchard, J. Mateo, M.F. Walsh, N. De Sarkar, W. Abida, H. Beltran, A. Garofalo, R. Gulati, S. Carreira, R. Eeles, et al., Inherited DNA-repair gene mutations in men with metastatic prostate cancer, N. Engl. J. Med. 375 (5) (2016) 443–453. https://10.1056/NEJMoa1603144.
[32] Y. Zong, J. Huang, D. Sankarasharma, T. Morikawa, M. Fukayama, J.I. Epstein, K.K. Chada, O.N. Witte, Stromal epigenetic dysregulation is sufficient to initiate mouse prostate cancer via paracrine Wnt signaling, Proc. Natl. Acad. Sci. U.S.A. 109 (50) (2012) E3395–E3404. https://10.1073/pnas.1217982109.
[33] Y. He, W. Xu, Y.T. Xiao, H. Huang, D. Gu, S. Ren, Targeting signaling pathways in prostate cancer: mechanisms and clinical trials, Signal Transduct. Targeted Ther. 7 (1) (2022) 198. https://10.1038/s41392-022-01042-7.
[34] B. Zhou, W. Lin, Y. Long, Y. Yang, H. Zhang, K. Wu, Q. Chu, Notch signaling pathway: architecture, disease, and therapeutics, Signal Transduct. Targeted Ther. 7 (1) (2022) 95. https://10.1038/s41392-022-00934-y.
[35] I. Mourkioti, A. Angelopoulou, K. Belogiannis, N. Lagopati, S. Potamianos, E. Kyrodimos, V. Gorgoulis, A. Papaspyropoulos, Interplay of developmental hippo-notch signaling pathways with the DNA damage response in prostate cancer, Cells 11 (15) (2022). https://10.3390/cells11152449.