Discover Oncology

Analysis

Identification of programmed cell death-related genes and construction of a prognostic model in oral squamous cell carcinoma using single-cell and transcriptome analysis

Yongheng Li[1] · Yang Yu[1] · Shaonan Hu[2] · Simin Li[2]

Received: 13 January 2025 / Accepted: 28 April 2025
Published online: 09 May 2025

Abstract

Background Oral squamous cell carcinoma (OSCC) is characterized by poor prognosis and high mortality. Understanding programmed cell death-related genes could provide valuable insights into disease progression and treatment strategies.
Methods RNA-sequencing data from 341 OSCC tumor tissues and 31 healthy samples were analyzed from TCGA database, with validation using 76 samples from GSE41613. Single-cell RNA sequencing data was obtained from GSE172577 (6 OSCC samples). Differentially expressed genes (DEGs) were identified and intersected with 1,254 programmed cell death-related genes. A protein–protein interaction network was constructed, and key modules were identified. Univariate Cox, LASSO, and multivariate Cox regression analyses were performed to build a prognostic model. Model performance was evaluated using Kaplan–Meier analysis, ROC curves, and nomogram validation.
Results The study identified 200 candidate genes from the intersection of DEGs and programmed cell death-related genes, which were further refined to 57 hub genes through PPI network analysis. A prognostic signature consisting of five genes (MET, GSDMB, KIT, PRKAG3, and CDKN2A) was established and validated. The model demonstrated good predictive performance in both training and validation cohorts (AUC > 0.6 for 1-, 2-, and 3-year survival). Single-cell analysis revealed that prognostic genes were predominantly expressed in stromal and epithelial cells. Cell communication analysis indicated strong interactions between stromal and epithelial cells.
Conclusions This study developed and validated a novel five-gene prognostic signature for OSCC based on programmed cell death-related genes. The model shows promising clinical application potential for risk stratification and personalized treatment of OSCC patients.

Keywords Oral squamous cell carcinoma · Programmed cell death · Prognostic model · Single-cell RNA sequencing · Tumor microenvironment

Senior author: Simin Li.

✉ Yongheng Li, 122304546@qq.com; Yang Yu, dr_yangyu@163.com; Shaonan Hu, shaonan_hu@smu.edu.cn; shaonan.hu@gmx.de; Simin Li, simin.li@qq.com; simin.li.dentist@gmail.com; siminli1988@smu.edu.cn | [1]Department of Stomatology, Qunli Branch, The First Affiliated Hospital of Harbin Medical University, 2075 Qunli Seventh Avenue, Harbin 150001, Heilongjiang Province, China. [2]Stomatological Hospital, School of Stomatology, Southern Medical University, 366 Jiangnan South Avenue, Haizhu District, Guangzhou 510280, Guangdong, China.

1 Introduction

Programmed cell death (PCD) represents a fundamental biological process that orchestrates the elimination of unnecessary or damaged cells, playing a pivotal role in tissue homeostasis and cancer development [1]. In oral squamous cell carcinoma (OSCC), multiple forms of PCD, including apoptosis, ferroptosis, necroptosis, pyroptosis, and autophagy, contribute to a complex network of cellular fate decisions that influence both tumor progression and treatment responses [2]. Recent advancements in PCD research have significantly expanded our understanding of its multifaceted forms and molecular mechanisms. Contemporary investigations are focused on developing therapeutic interventions targeting apoptotic pathways, particularly in synergistic combinations with immune checkpoint inhibitors [3, 4]. Autophagy-dependent cell death exhibits dichotomous functionality, functioning as a tumor suppressor through elimination of damaged constituents while simultaneously serving as a survival mechanism under conditions of cellular stress. The modulation of autophagic processes is currently being investigated as a potential therapeutic strategy for oncological pathologies [5, 6]. Necroptosis has emerged as a focal point of innovative research aimed at harnessing its pathways to potentiate cancer immunotherapeutic efficacy [7]. Pyroptosis has recently demonstrated significant capacity to augment anti-tumor immune responses when pharmacologically induced [8]. Ferroptosis is currently the subject of pioneering research exploring its inducers as novel therapeutic modalities for malignancies resistant to conventional treatments [7]. The identification of emergent PCD forms, such as cuproptosis, has substantially expanded the repertoire of cell death mechanisms [9], while state-of-the-art synthetic biology approaches are being developed to precisely manipulate specific death processes with therapeutic intent [8]. The current research paradigm is increasingly oriented toward elucidating the potential applications of PCD in enhancing immunotherapeutic interventions, representing a promising frontier in oncological research [10].

The dysregulation of apoptotic pathways, particularly through alterations in Bcl-2 family proteins, has been well-documented in oral tumors, leading to uncontrolled cell proliferation and therapy resistance [2]. Recent research has highlighted the emerging role of ferroptosis, an iron-dependent form of cell death, as a promising therapeutic target in OSCC through its unique mechanism of inducing oxidative stress-mediated cell death [11]. Additionally, necroptosis and pyroptosis, mediated by RIPK1/RIPK3 and inflammatory caspases respectively, have been shown to significantly influence the tumor microenvironment through their inflammatory effects [12]. The dual role of autophagy in OSCC has attracted considerable attention, as it can either promote cancer cell survival or contribute to cell death depending on the cellular context [13]. Moreover, the intricate crosstalk between different PCD pathways has emerged as a critical factor in determining treatment outcomes, with recent studies revealing complex interactions between these cell death mechanisms [14]. The modulation of specific PCD pathways, such as targeting GPX4 in ferroptosis or anti-apoptotic proteins in apoptosis, represents a promising avenue for therapeutic intervention in OSCC [15]. Understanding these diverse PCD mechanisms and their regulatory networks has become increasingly crucial for developing more effective treatment strategies for OSCC patients.

Recent studies have demonstrated that the complex interplay of PCD mechanisms significantly influences the prognosis and treatment outcomes in OSCC. The dysregulation of PCD pathways, particularly ferroptosis, apoptosis, and autophagy, has been shown to contribute to tumor progression and therapeutic resistance [2, 16]. While individual PCD-related genes have been investigated, a comprehensive understanding of their collective prognostic value remains limited. The emergence of ferroptosis as a distinct form of PCD has particularly attracted attention, with studies revealing its potential as both a prognostic indicator and therapeutic target in cancer treatment [17]. Moreover, the intricate crosstalk between different PCD pathways, mediated by various molecular factors including non-coding RNAs, suggests that a single-gene approach may be insufficient to accurately predict patient outcomes [14]. The complex nature of PCD regulation in OSCC, involving multiple interconnected pathways and mechanisms, necessitates the development of a more comprehensive prognostic model that can capture these multilayered interactions [15]. Such a model, incorporating multiple PCD-related genes, could provide more accurate risk stratification and guide personalized treatment decisions. Furthermore, the identification of key PCD-related genes and their prognostic significance could reveal novel therapeutic targets and strategies, particularly in the context of treatment resistance and tumor recurrence. Therefore, establishing a robust prognostic model based on PCD-related genes could not only enhance our ability to predict patient outcomes but also contribute to the development of more effective therapeutic approaches in OSCC.

In this study, we aimed to establish a novel prognostic model based on PCD-related genes in OSCC through integrated analysis of transcriptomic data and single-cell RNA sequencing. By combining differential expression

analysis, protein–protein interaction network construction, and machine learning approaches, we sought to identify key PCD-related genes and evaluate their prognostic value. Furthermore, we investigated the cellular distribution and potential mechanisms of these genes through single-cell analysis, aiming to provide new perspectives for personalized treatment strategies in OSCC patients. This comprehensive analysis not only enhances our understanding of PCD-related mechanisms in OSCC but also offers potential therapeutic targets for clinical intervention.

2 Methods

2.1 Data collection and processing

RNA sequencing data from 341 OSCC tumor tissues and 31 healthy samples were obtained from The Cancer Genome Atlas (TCGA-HNSC) database [18]. For validation purposes, the GSE41613 dataset containing 76 OSCC samples with confirmed cancer-specific mortality was downloaded from the Gene Expression Omnibus (GEO) database [19]. Single-cell RNA sequencing data were acquired from the GSE172577 dataset (GPL24676), comprising six OSCC tumor tissue samples. A comprehensive list of 1,254 programmed cell death-related genes was obtained.

We utilized the TCGA-HNSC level 3 HTSeq-FPKM-UQ normalized RNA-sequencing data from the GDC Data Portal (release 42.0). For data preprocessing, we converted FPKM-UQ values to TPM (Transcripts Per Million) to improve comparability across samples. Genes with low expression (TPM < 1 in more than 50% of samples) were filtered out. To address potential batch effects, we applied the ComBat algorithm from the sva R package, using sequencing center and plate as covariates. Data normalization included $\log_2$ transformation ($\log_2$[TPM + 1]) for downstream analysis. Quality control included assessment of sample correlation and principal component analysis to identify outliers, which were subsequently removed before differential expression analysis.

For the GSE41613 dataset, we downloaded the raw CEL files and processed them using the affy package in R. Background correction was performed using the robust multi-array average (RMA) method, followed by quantile normalization and $\log_2$ transformation. Probe sets were mapped to gene symbols using the corresponding annotation package. When multiple probe sets mapped to the same gene, we retained the probe set with the highest mean expression value across all samples. For sample inclusion criteria, we excluded samples with incomplete survival information or follow-up less than 30 days. In the case of GSE41613, we specifically selected samples where death was confirmed to be caused by oral cancer, resulting in the final validation cohort of 76 patients. Missing values in the clinical data were handled using multiple imputation for variables with less than 20% missing data. Variables with more than 20% missing data were excluded from multivariate analyses. For gene expression data, no imputation was performed; instead, genes with excessive missing values were filtered out during the low-expression filtering step. Potential outliers were identified using Cook's distance in the regression models, and sensitivity analyses were performed to assess their impact on the results.

2.2 Identification of differentially expressed genes and hub genes

Differential expression analysis was performed using the limma package in R. Genes with $|\log_2 FC| \geq 1$ and p-value <0.05 were considered significantly differentially expressed. The intersection between differentially expressed genes and programmed cell death-related genes was identified. A protein–protein interaction (PPI) network was constructed using the STRING database (confidence score >0.4). Hub genes were identified through the analysis of key modules using the MCODE plugin in Cytoscape. Functional enrichment analysis of hub genes was conducted using the clusterProfiler R package, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses.

2.3 Construction and validation of the prognostic model

Univariate Cox regression analysis was initially performed to identify survival-associated genes ($p < 0.01$). The proportional hazards assumption was verified, and genes meeting this assumption were retained for further analysis. LASSO regression was then applied to select key prognostic genes and construct the risk model. To determine the optimal regularization parameter lambda, we performed tenfold cross-validation, where the dataset was randomly partitioned into 10 equal-sized subsets. Nine subsets were used for training while the remaining subset served as validation data. This process was repeated 10 times, ensuring each subset was used exactly once for validation. The lambda value that minimized

the cross-validation error (lambda.min) was selected for the final model. Cross-validation standard errors were also calculated to assess model stability. Multivariate Cox regression analysis was subsequently performed to establish the final prognostic signature. The risk score for each patient was calculated using the formula: Risk score = Σ(coef(gene_i) × expr(gene_i)), where coef represents the regression coefficient and expr represents the gene expression level.

2.4 Gene–gene interaction (GGI) network analysis

To investigate the functional model and interacting proteins of the key prognostic genes, we constructed a gene–gene interaction network using GeneMANIA (http://genemania.org). The five prognostic genes identified from our model (MET, GSDMB, KIT, PRKAG3, and CDKN2 A) were used as seed genes for network construction. GeneMANIA integrates multiple sources of evidence including physical interactions, co-expression patterns, predicted interactions, co-localization, genetic interactions, pathway connections, and shared protein domains to build comprehensive interaction networks. For visualization purposes, we selected the top 20 genes that showed the strongest interactions with our five key genes. The network was configured to display seven types of interactions: physical interactions, co-expression, predicted interactions, co-localization, genetic interactions, pathway connections, and shared protein domains. Additionally, the top seven enriched functions associated with the network were identified and color-coded to illustrate the functional relevance of these gene interactions.

2.5 Validation and performance assessment

The prognostic model's performance was evaluated through Kaplan–Meier survival analysis and time-dependent ROC curves for 1-, 2-, and 3-year survival predictions with the 'survival' package in R. A nomogram incorporating independent prognostic factors was constructed using the rms R package [20]. Calibration curves and decision curve analysis (DCA) [21] were generated to assess the nomogram's predictive accuracy. Calibration curves were generated to evaluate the agreement between nomogram-predicted probabilities and actual observed outcomes for 1-, 2-, and 3-year survival. Patients were grouped based on their predicted probabilities, and the mean predicted probability in each group was plotted against the observed outcome frequency using the Kaplan–Meier method. The calibration curves were fitted using a flexible spline function with 95% confidence intervals calculated through bootstrap resampling with 1,000 iterations. The closer the calibration curve to the 45-degree diagonal line, the better the agreement between prediction and observation. To assess the clinical utility of the nomogram, DCA was performed using the 'ggDCA' package. DCA quantifies the net benefit of using the prediction model across a range of threshold probabilities, compared to treating all patients or treating none. Net benefit was calculated as the difference between the true-positive rate and the weighted false-positive rate, with the weight equal to the odds of the threshold probability. The higher the curve on the DCA plot, the greater the net benefit of using that model for clinical decision-making. The model's performance was validated using the independent GSE41613 dataset.

2.6 Immune cell infiltration analysis

Immune cell infiltration analysis was conducted using the CIBERSORTx algorithm [22] with the leukocyte signature matrix (LM22), which distinguishes 22 immune cell subtypes. Bulk RNA-seq data from TCGA-HNSC was uploaded to the CIBERSORTx web portal (https://cibersortx.stanford.edu/) and analyzed with 1,000 permutations and quantile normalization enabled for the reference dataset. Samples with a CIBERSORTx p-value <0.05 were considered to have reliable immune cell proportion estimates and were included in subsequent analyses. The relative proportions of different immune cell types between high- and low-risk groups were compared using the Wilcoxon rank-sum test, with Benjamini–Hochberg correction for multiple testing. Correlation between risk scores and immune cell proportions was assessed using Spearman's rank correlation analysis.

To investigate the relationship between immune-regulatory genes and our risk model, we performed Spearman's rank correlation analysis between risk scores and the expression of immune-regulatory factors. The immune-regulatory genes analyzed in our study encompass six major functional categories: HLA family (HLA-A, HLA-B, HLA-C, etc.), adhesion molecules (ICAM1, ITGB2, etc.), immune checkpoints (CD274, CD276, etc.), cytokines/chemokines (CCL5, CXCL9, IL10, etc.), immunomodulatory enzymes (ARG1, ENTPD1, etc.), and immune cell receptors (BTLA, CTLA4, PDCD1, etc.). For each immune-regulatory gene, Spearman's correlation coefficient was calculated against the risk score derived from

our five-gene signature. Statistical significance was determined with *p*-values adjusted for multiple testing using the Benjamini–Hochberg method.

2.7 Single-cell RNA sequencing analysis

Single-cell RNA sequencing data were processed using the Seurat package (version 5.0.0). Quality control parameters were set as follows: nFeature_RNA between 200 and 4000, nCount_RNA <20,000, and percent.mt <15%. Data normalization was performed using the LogNormalize method, and variable features were identified using the vst method. Principal component analysis was conducted, followed by harmony-based integration of samples to correct for batch effects. Specifically, we used the harmony algorithm [23] implemented in the harmonize function of the Seurat package, with the parameter 'reduction ="pca"' and using sample identity as the batch variable. This approach allows for the correction of technical variation between samples while preserving biological heterogeneity. The harmony integration was performed on the first 30 principal components before downstream clustering and visualization. The effectiveness of batch correction was evaluated by examining the mixing of cells from different samples in the reduced dimensional space.

Cell clustering was performed using FindNeighbors and FindClusters functions with a resolution of 0.5. For cell type annotation, we used the ACT database (accessed from http://xteam.xbio.top/ACT/ in January 2024) based on the top 30 marker genes for each cluster. Marker genes were identified using the FindAllMarkers function with the following parameters: min.pct = 0.2 (minimum percentage of cells expressing the gene in either of the two compared groups), logfc.threshold = 0.25 (minimum log fold-change), and test.use ="roc" (using the ROC test for marker detection). Only genes with an adjusted *p*-value <0.05 were considered as significant markers. The top 30 marker genes ranked by average log fold-change were used for cell type annotation through the ACT database's automated annotation pipeline. Cell–cell communication analysis was performed using the CellChat R package to evaluate potential ligand-receptor interactions between different cell types. Default parameters were used for the analysis, with a *p*-value threshold of 0.05 for significant interactions.

2.8 Somatic mutation data analysis

To explore the somatic mutation profiles between high and low-risk groups, we analyzed somatic mutation data from TCGA database. The mutation data was processed and visualized using the "maftools" package (version 2.12.0) in R. We first categorized patients into high and low-risk groups based on the previously established risk score model. Mutation annotation format (MAF) files were downloaded from the TCGA database, containing detailed information about each somatic mutation. The "read.maf" function was used to import and summarize the mutation data. We compared mutation frequencies, spectra, and significantly mutated genes between the high and low-risk groups. In our analysis, we classified mutations into several types including Missense_Mutation, Nonsense_Mutation, Frame_Shift_Del, Frame_Shift_Ins, Splice_Site, In_Frame_Del, Translation_Start_Site, and Multi_Hit. Oncoplot visualizations were generated to display the distribution of these mutation types across samples for each risk group. Statistical significance of differential mutation rates between groups was assessed using Fisher's exact test with Benjamini–Hochberg correction for multiple comparisons. Genes with a significantly different mutation frequency between risk groups (adjusted $p < 0.05$) were identified as differentially mutated genes.

2.9 Drug sensitivity analysis

To assess the potential therapeutic implications of the prognostic signature, we conducted drug sensitivity analysis using the oncoPredict package in R. The Genomics of Drug Sensitivity in Cancer (GDSC) database v2 was used as the training dataset, which comprises expression profiles of 17,419 genes across 805 cell lines and drug response data for 198 compounds. We constructed a predictive model based on this training set to estimate the half maximal inhibitory concentration (IC50) values of various drugs for each OSCC sample in our cohort. The IC50 values, representing the required drug concentration to achieve 50% inhibition of cancer cell growth, were calculated for all available compounds. Statistical comparisons of IC50 values between high- and low-risk groups were performed using the Wilcoxon rank-sum test, with Benjamini–Hochberg correction for multiple testing. Additionally, Spearman's correlation analysis was conducted to assess the association between risk scores and drug sensitivity.

3 Results

3.1 Identification and visualization of differentially expressed genes between oral squamous cell carcinoma and normal tissues

The differential expression analysis identified 4,300 differentially expressed genes (DEGs) between OSCC tumor tissues and normal controls, including 1,930 upregulated and 2,370 downregulated genes (|log2 FC| ≥ 1 and $p < 0.05$). As shown in Fig. 1, the volcano plot (Fig. 1A) displays the distribution of DEGs, with significantly upregulated genes (red dots) including MAGEA family members and CA9, while downregulated genes (blue dots) include BPIFA2, HTN1, HTN3, and SMR3B. The heatmap visualization (Fig. 1B) demonstrates distinct expression patterns between tumor and normal samples, with a clear clustering of upregulated and downregulated genes. A circular heatmap (Fig. 1C) further illustrates the expression

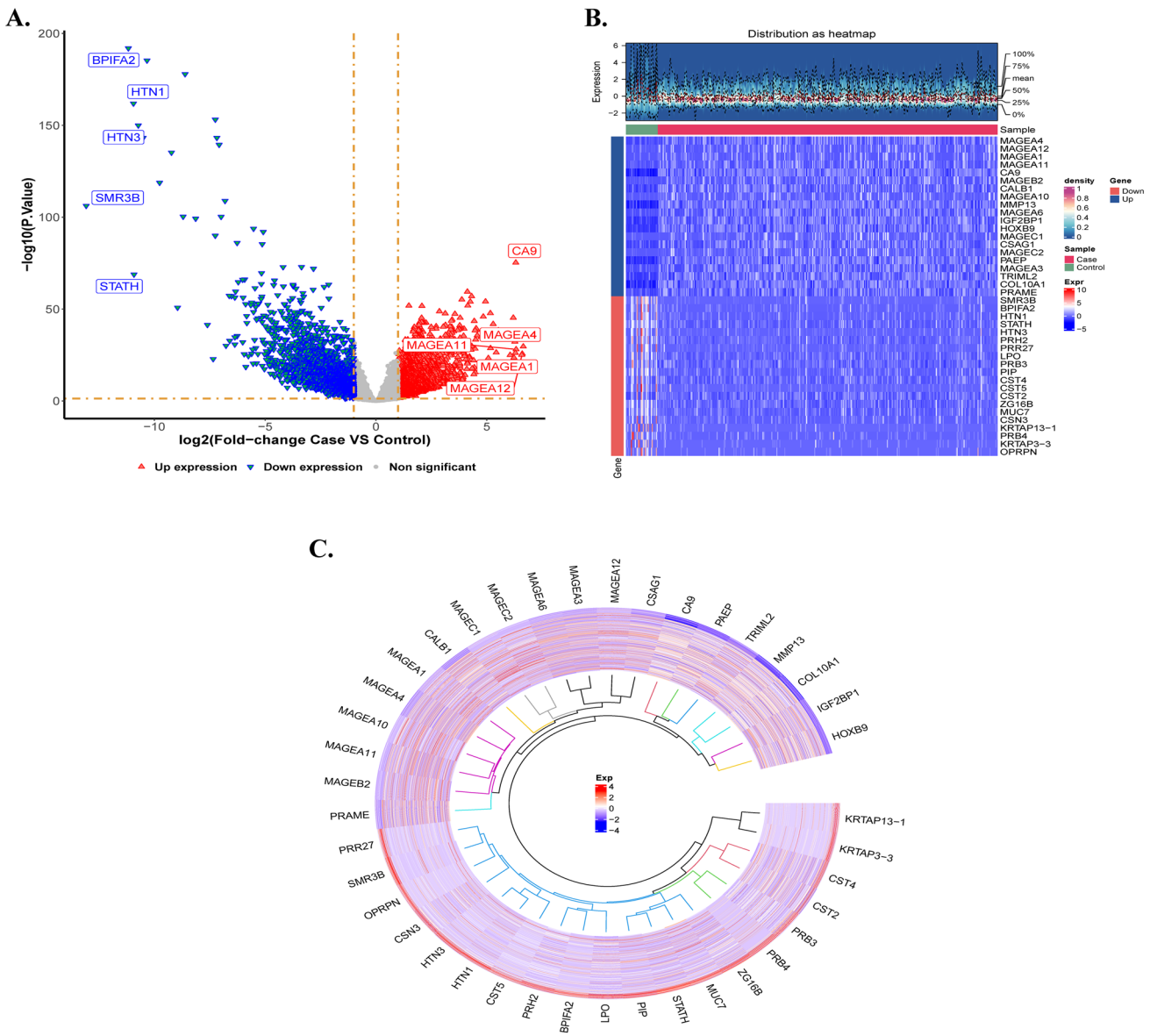


Fig. 1 Differential gene expression analysis in OSCC and normal tissues. **A** Volcano plot showing differentially expressed genes between OSCC and normal tissues. Red dots represent upregulated genes and blue dots represent downregulated genes (|log2 FC| ≥ 1, $p < 0.05$). **B** Heatmap showing the hierarchical clustering of differentially expressed genes. **C** Circular visualization of top differential genes' expression patterns across all samples

profiles of the top differentially expressed genes across all samples, providing an alternative visualization of the gene expression patterns that distinguish OSCC from normal tissues.

3.2 Functional enrichment analysis of programmed cell death-related genes in OSCC

The comprehensive analysis of programmed cell death (PCD)-related genes in OSCC revealed multiple significant findings as shown in Fig. 2. The Venn diagram (Fig. 2A) identified 200 overlapping genes (3.7%) between differentially expressed genes (DEGs) and PCD-related genes. The PPI network analysis (Fig. 2B) visualized the complex interactions among these candidate genes. GO analysis (Fig. 2C) demonstrated enrichment in biological processes related to apoptotic signaling pathways, with the results displayed in both bar plot and treemap formats (Fig. 2D). The treemap visualization highlighted the hierarchical relationship of GO terms across biological processes (BP), cellular components (CC), and molecular functions (MF). KEGG pathway analysis (Fig. 2E) identified significantly enriched pathways displayed in a bubble plot, where the size of each bubble represents the number of genes and the color indicates the adjusted *p*-value. The circular network diagram (Fig. 2F) further illustrated the relationships between genes and key pathways, particularly highlighting cytokine-cytokine receptor interaction, rheumatoid arthritis, lysosome, ferroptosis, and IL-17 signaling pathways as the most significantly enriched KEGG pathways. The analysis revealed that these PCD-related genes are primarily involved in immune response, inflammatory processes, and cell death pathways in OSCC.

3.3 Construction and validation of a prognostic risk model based on PCD-related genes in OSCC

Figure 3 presents the comprehensive development and validation of a prognostic risk model. Through univariate Cox regression analysis (Fig. 3A), seven genes were identified as significantly associated with survival ($p < 0.05$). LASSO regression analysis (Fig. 3B, C) was performed to prevent overfitting, with the optimal lambda value of 0.019 selected. Multivariate Cox regression analysis (Fig. 3D) further refined the model to five independent prognostic genes: MET (HR = 1.3), GSDMB (HR = 1.2), KIT (HR = 0.74), PRKAG3 (HR = 1.4), and CDKN2 A (HR = 0.93).

To understand the functional context of the five prognostic genes (MET, GSDMB, KIT, PRKAG3, and CDKN2A), we constructed a GGI network as shown in Fig. 3E. The network revealed extensive interactions among our five key genes and their top 20 interacting partners. Functional enrichment analysis of the network revealed seven predominant functions, with peptidyl-tyrosine phosphorylation and protein tyrosine kinase activity being the most significant. Other important functions included regulation of peptidyl-tyrosine phosphorylation, cellular response to insulin stimulus, insulin receptor signaling pathway, and receptor tyrosine kinase binding.

The risk score distribution and survival status plots for both training (Fig. 3F) and validation (Fig. 3I) cohorts demonstrated the model's discriminative ability. Kaplan–Meier survival analyses revealed significant survival differences between high- and low-risk groups in both the training (Fig. 3G, $p < 0.0001$) and validation (Fig. 3J, $p = 0.034$) cohorts. Time-dependent ROC curve analysis showed the model's predictive accuracy at 1-, 2-, and 3-year intervals in the training cohort (Fig. 3H, AUC = 0.631–0.667) and validation cohort (Fig. 3K, AUC = 0.61–0.67), confirming the model's robust prognostic value.

3.4 Development and validation of a nomogram for predicting OSCC patient survival

Figure 4 presents a comprehensive evaluation of independent prognostic factors and the construction of a predictive nomogram. Univariate Cox regression analysis (Fig. 4A) and multivariate Cox regression analysis (Fig. 4B) identified the risk score as an independent prognostic factor (HR = 2.66, 95% CI: 1.8–3.95, $p < 0.001$). Based on the five identified genes (MET, GSDMB, KIT, PRKAG3, and CDKN2 A), a nomogram was constructed (Fig. 4C) to predict 1-, 2-, and 3-year overall survival probabilities in OSCC patients. The calibration curves (Fig. 4D) demonstrated good agreement between nomogram predictions and actual observations (C-index = 0.62). Decision curve analysis (DCA) for 1-year (Fig. 4E), 2-year (Fig. 4F), and 3-year (Fig. 4G) survival predictions showed that the nomogram offered superior net clinical benefit compared to individual genes across a wide range of threshold probabilities. The DCA curves indicate that the nomogram-based predictions outperform the "treat-all" or "treat-none" strategies, validating its clinical utility in guiding personalized treatment decisions for OSCC patients.

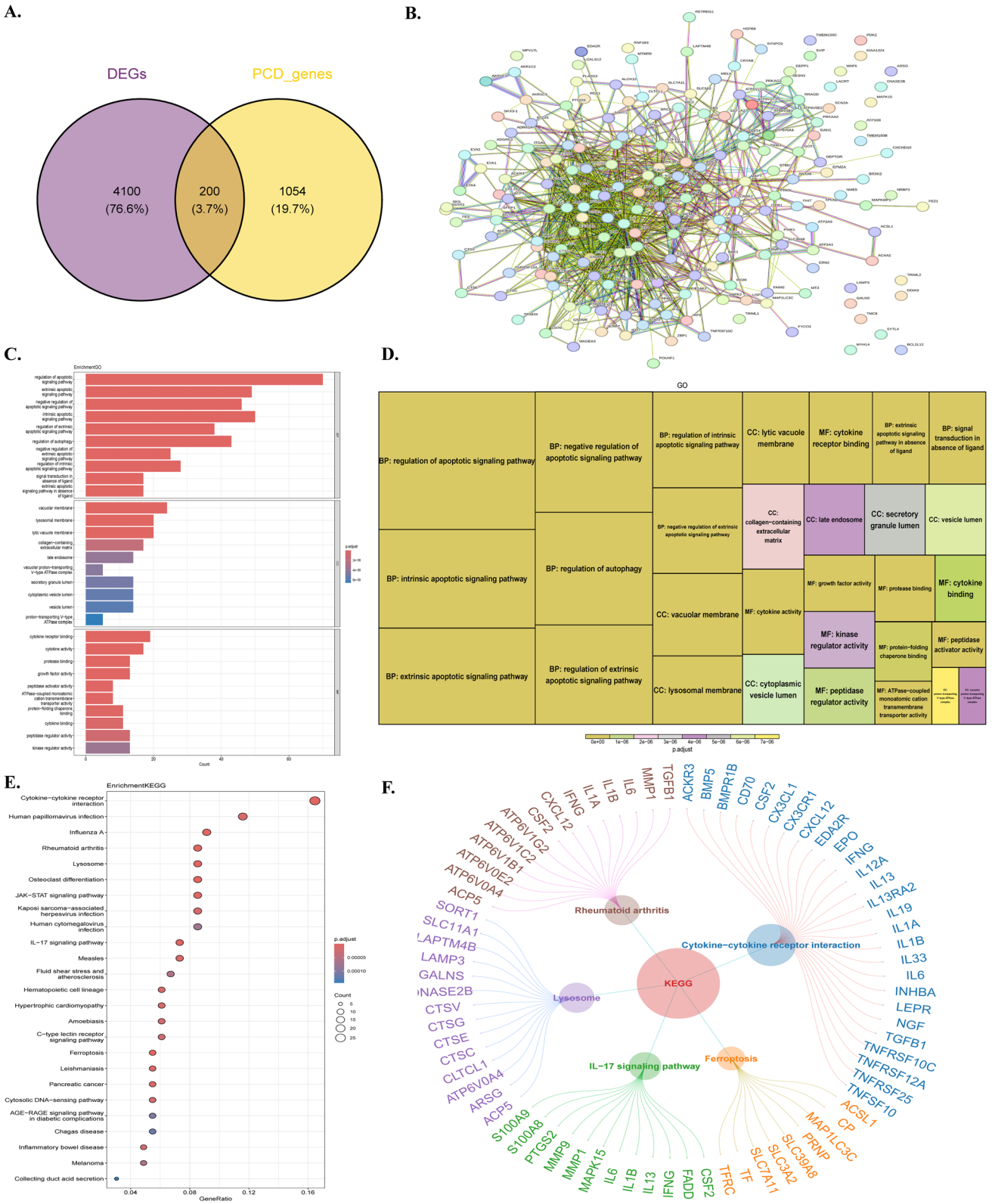


Fig. 2 Identification and functional enrichment analysis of candidate genes. **A** Venn diagram showing the overlap between DEGs and PCD-related genes. **B** Protein–protein interaction network of candidate genes. **C** Bar plot of GO enrichment analysis for biological process, cellular component and molecular function. **D** Treemap visualization of GO enrichment results. **E** Bubble plot showing KEGG pathway enrichment analysis. **F** Circular network diagram illustrating relationships between genes and enriched KEGG pathways

3.5 GSEA reveals distinct pathway enrichment patterns for prognostic genes

Figure 5 presents the Gene Set Enrichment Analysis (GSEA) results for the five prognostic genes in OSCC. For MET (Fig. 5A), significant enrichment was observed in oxidative phosphorylation and focal adhesion pathways (positive correlation), while showing negative enrichment in ribosome and Parkinson's disease pathways. GSDMB (Fig. 5B) showed strong positive enrichment in multiple immune-related pathways, including antigen processing and presentation, graft versus host disease, and allograft rejection. KIT (Fig. 5C) demonstrated negative enrichment across multiple pathways, particularly in ribosome, spliceosome, and oxidative phosphorylation pathways. PRKAG3 (Fig. 5D) exhibited notable enrichment in cardiomyopathy-related pathways and cell cycle regulation, with both positive and negative correlations observed. CDKN2A (Fig. 5E) showed significant positive enrichment in DNA-related processes, including DNA replication, cell cycle, and various DNA repair pathways, particularly the spliceosome pathway. These enrichment patterns suggest that these prognostic genes are involved in diverse biological processes ranging from metabolism and immune response to cell cycle regulation and DNA repair in OSCC development and progression.

3.6 Analysis of immune cell infiltration patterns and their correlation with prognostic genes in OSCC

Figure 6 presents a comprehensive analysis of the immune cell landscape in OSCC. The box plot (Fig. 6A) shows the distribution of 22 immune cell types between high- and low-risk groups, with several cell types showing significant differences ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$). The stacked bar plot (Fig. 6B) visualizes the proportion of different immune cell types in each sample, clearly demonstrating the heterogeneity of immune cell composition across high- and low-risk groups. The heatmap (Fig. 6C) illustrates the correlation between the five prognostic genes (MET, GSDMB, KIT, PRKAG3, and CDKN2A) and various immune cell types, showing both positive (red) and negative (blue) correlations. Notably, significant correlations were observed between these genes and specific immune cell populations, particularly with T cells, macrophages, and dendritic cells. The color intensities and significance levels (indicated by asterisks) reflect the strength and statistical significance of these correlations, providing insights into the potential immune regulatory mechanisms associated with these prognostic genes in OSCC. As shown in Fig. 6D, immune-regulatory gene analysis revealed significant positive correlations ($p \leq 0.0001$) between risk scores and several genes including HLA-C, MICB, CD276, and TGFB1, suggesting enhanced antigen presentation and immunosuppressive signaling in high-risk patients. Meanwhile, strong negative correlations ($p \leq 0.0001$) were observed with multiple T-cell activation and co-stimulatory molecules including CD28, ICOSLG, CD40LG, IL2, TNFSF4, ENTPD1, BTLA, CD27, and TNFRSF9, indicating potential impairment of T-cell function and reduced immune surveillance in patients with elevated risk scores.

3.7 Single-cell RNA sequencing reveals cellular heterogeneity and expression patterns of prognostic genes in OSCC

Figure 7 presents a comprehensive single-cell RNA sequencing analysis of OSCC samples. Quality control metrics (Fig. 7A) show the distribution of gene counts (nFeature_RNA), RNA molecules (nCount_RNA), and mitochondrial gene percentage (percent.mt) across six samples. Variable gene analysis (Fig. 7B) identified highly variable genes, with key genes highlighted. Principal Component Analysis (PCA) revealed distinct cellular clusters (Fig. 7C), with the elbow plot (Fig. 7D) helping determine optimal principal components for downstream analysis. UMAP visualization in Fig. 7E shows distinct cellular clustering from six OSCC samples, while Fig. 7F displays the same projection after annotation, revealing 13 distinct cell populations including epithelial, immune, and stromal components that comprise the complex OSCC tumor microenvironment. The dot plot (Fig. 7G) visualizes the expression patterns of the five prognostic genes (MET, GSDMB, KIT, PRKAG3, CDKN2 A) across different cell types, with dot size representing the percentage of expressing cells and color intensity indicating expression level. Our cell–cell communication analysis using the CellChat package revealed important intercellular signaling patterns within the OSCC tumor microenvironment, as illustrated in Fig. 7H. The analysis demonstrated that stromal cells and epithelial cells exhibited particularly high interaction intensity, indicating strong cellular communication between these two populations.

3.8 Somatic mutation analysis results

To investigate the genomic alterations associated with different risk groups, we analyzed somatic mutation profiles in both high and low-risk OSCC patients. As shown in Fig. 8, the mutation landscape exhibited distinct patterns between

Fig. 3 Construction and validation of the programmed cell death-related prognostic model for OSCC. **A** Forest plot of univariate Cox regression analysis identifying survival-associated genes ($p < 0.05$). **B** LASSO coefficient profiles of the candidate prognostic genes. Each curve represents a gene, with coefficient values changing along the log(lambda) parameter. **C** Cross-validation plot for selecting the optimal lambda value in LASSO regression. **D** Forest plot of multivariate Cox regression analysis showing the five independent prognostic genes (MET, GSDMB, KIT, PRKAG3, and CDKN2 A) with their respective hazard ratios and 95% confidence intervals. **E** Gene–gene interaction network showing the functional associations between the five prognostic genes and their top interacting partners. **F** Risk score distribution (upper panel) and patient survival status (lower panel) in the TCGA training cohort. Patients are arranged by increasing risk score from left to right. **G** Kaplan–Meier survival curve comparing high- and low-risk groups in the TCGA training cohort ($p < 0.0001$). **H** Time-dependent ROC curves showing the predictive performance of the risk model for 1-, 2-, and 3-year survival in the TCGA training cohort (AUC = 0.631–0.667). **I** Risk score distribution (upper panel) and patient survival status (lower panel) in the GSE41613 validation cohort. **J** Kaplan–Meier survival curve comparing high- and low-risk groups in the GSE41613 validation cohort ($p = 0.034$). **K** Time-dependent ROC curves showing the predictive performance of the risk model for 1-, 2-, and 3-year survival in the GSE41613 validation cohort (AUC = 0.61–0.67)

the two risk groups. In the low-risk group (Fig. 8A), 152 out of 164 samples (92.68%) displayed genomic alterations. TP53 was the most frequently mutated gene (63%), followed by TTN (32%), CDKN2A (27%), FAT1 (24%), PIK3 CA (20%), and NOTCH1 (19%). Other recurrently mutated genes included CASP8 (16%), MUC16 (15%), KMT2D (14%), SYNE1 (12%), CSMD3 (12%), LRP1B (12%), FLG (11%), DNAH5 (10%), PCLO (10%), DOP1 A (9%), EP300 (9%), FBXW7 (9%), PCDH11X (9%), and SYNE2 (9%). In the high-risk group (Fig. 8B), a higher mutation rate was observed with 160 out of 164 samples (97.56%) harboring mutations. TP53 mutations were markedly more prevalent in the high-risk group (79%) compared to the low-risk group. TTN showed the same mutation frequency (32%) as in the low-risk group. Other frequently mutated genes in the high-risk group included FAT1 (24%), CDKN2A (20%), NOTCH1 (20%), SYNE1 (16%), CASP8 (15%), CSMD3 (15%), HUWE1 (15%), MUC16 (14%), DNAH5 (12%), PCLO (12%), PIK3 CA (12%), AHNAK (12%), LRP1B (12%), KMT2D (10%), PLEC (10%), FAT3 (9%), XIRP2 (9%), and HRAS (9%). Notably, we observed a significant difference in TP53 mutation frequency between high-risk (79%) and low-risk (63%) groups. Additionally, the mutation profiles of several genes varied between the two groups, with genes like HUWE1, AHNAK, PLEC, and HRAS showing higher mutation frequencies in the high-risk group, while CDKN2A showed higher mutation frequency in the low-risk group (27 vs. 20%). These findings suggest that distinct mutational processes might be operative in different risk groups, potentially contributing to their differential clinical outcomes.

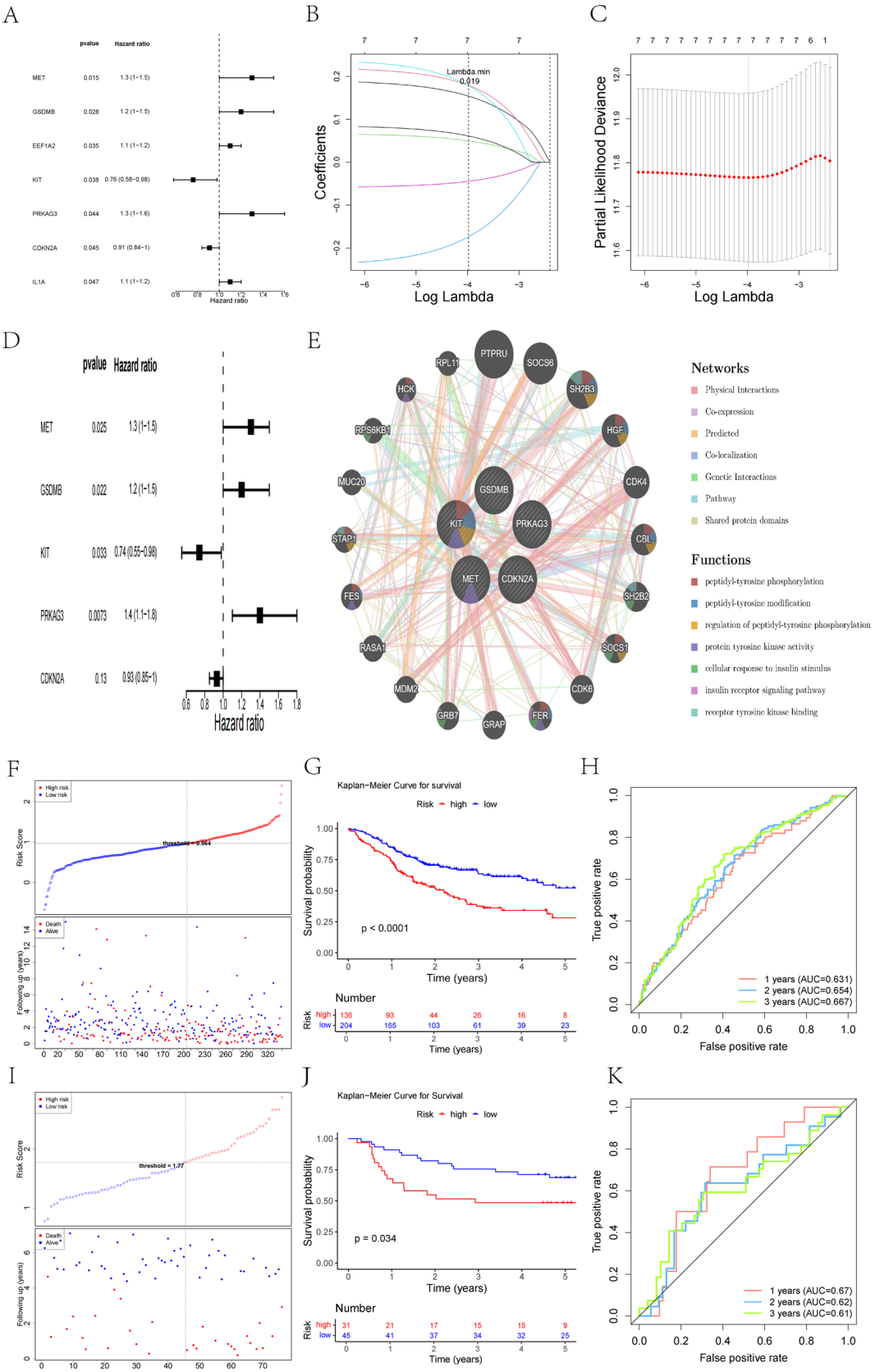
3.9 Drug sensitivity analysis in relation to risk score

Drug sensitivity analysis revealed significant correlations between risk scores and responses to specific therapeutic compounds. As shown in Fig. 8C, the heatmap correlation analysis demonstrated that risk scores exhibited distinct correlation patterns with IC50 values of multiple drugs. Notably, we identified three drugs that showed strong positive correlations with risk scores (correlation coefficients >0.3, $p < 0.05$). Figure 8D–F illustrate the detailed correlation analyses for three key drugs with risk scores: GSK2699G2 A_1192 (R = 0.32, $p = 1.6e{-}09$), Venetoclax_1909 (R = 0.33, $p = 7.4e{-}10$), and ABT737_1910 (R = 0.31, $p = 1.1e{-}08$). Both scatter plots and box plots demonstrate that patients in the high-risk group (shown in red) exhibited significantly lower sensitivity to these three drugs compared to those in the low-risk group (shown in blue), as evidenced by higher IC50 values in the high-risk group.

4 Discussion

In this study, we conducted a comprehensive analysis of programmed cell death-related genes in oral squamous cell carcinoma and established a robust prognostic model. Our analysis identified a novel five-gene signature (MET, GSDMB, KIT, PRKAG3, and CDKN2A) that effectively stratified OSCC patients into high- and low-risk groups with significantly different survival outcomes. Through single-cell RNA sequencing analysis, we discovered that these prognostic genes were predominantly expressed in stromal and epithelial cells, with strong cell–cell communication patterns observed between these populations. Furthermore, the prognostic signature was found to be closely associated with distinct immune cell infiltration patterns, suggesting its potential utility in guiding personalized treatment strategies.

The intricate roles of the five genes MET, GSDMB, KIT, PRKAG3, and CDKN2A in OSCC are crucial for understanding their implications in cancer prognosis and therapy. While specific references to MET gene in OSCC prognosis are not deeply highlighted in recent searches, generally, MET is known for its involvement in cell growth, motility, and differentiation, often implicated in various cancers for promoting metastatic behavior [24]. The MET gene encodes a protein tyrosine kinase that can activate a number of downstream signaling pathways affecting cell division and movement, thus


A
pvalue
Hazard ratio
MET 0.015 1.3 (1−1.5)
GSDMB 0.028 1.2 (1−1.5)
EEF1A2 0.035 1.1 (1−1.2)
KIT 0.038 0.76 (0.58−0.98)
PRKAG3 0.044 1.3 (1−1.6)
CDKN2A 0.045 0.91 (0.84−1)
IL1A 0.047 1.1 (1−1.2)
Hazard ratio
B
Lambda.min
0.019
Coefficients
Log Lambda
C
Partial Likelihood Deviance
Log Lambda
D
pvalue Hazard ratio
MET 0.025 1.3 (1−1.5)
GSDMB 0.022 1.2 (1−1.5)
KIT 0.033 0.74 (0.55−0.98)
PRKAG3 0.0073 1.4 (1.1−1.8)
CDKN2A 0.13 0.93 (0.85−1)
Hazard ratio
E
PTPRU
SOCS6
RPL11
HCK
SH2B3
RPS6KB1
HGF
MUC20
GSDMB
CDK4
KIT
PRKAG3
STAP1
CBL
MET
CDKN2A
FES
SH2B2
RASA1
SOCS1
MDM2
CDK6
GRB7
GRAP
FER
Networks
Physical Interactions
Co-expression
Predicted
Co-localization
Genetic Interactions
Pathway
Shared protein domains
Functions
peptidyl-tyrosine phosphorylation
peptidyl-tyrosine modification
regulation of peptidyl-tyrosine phosphorylation
protein tyrosine kinase activity
cellular response to insulin stimulus
insulin receptor signaling pathway
receptor tyrosine kinase binding
F
High risk
Low risk
Risk Score
threshold = 0.964
Death
Alive
Following up (years)
G
Kaplan−Meier Curve for survival
Risk high low
Survival probability
p < 0.0001
Time (years)
Number
Risk high 136 93 44 26 16 8
low 204 165 103 61 39 23
H
True positive rate
False positive rate
1 years (AUC=0.631)
2 years (AUC=0.654)
3 years (AUC=0.667)
I
High risk
Low risk
Risk Score
threshold = 1.77
Death
Alive
Following up (years)
J
Kaplan−Meier Curve for Survival
Risk high low
Survival probability
p = 0.034
Time (years)
Number
Risk high 31 21 17 15 15 9
low 45 41 37 34 32 25
K
True positive rate
False positive rate
1 years (AUC=0.67)
2 years (AUC=0.62)
3 years (AUC=0.61)

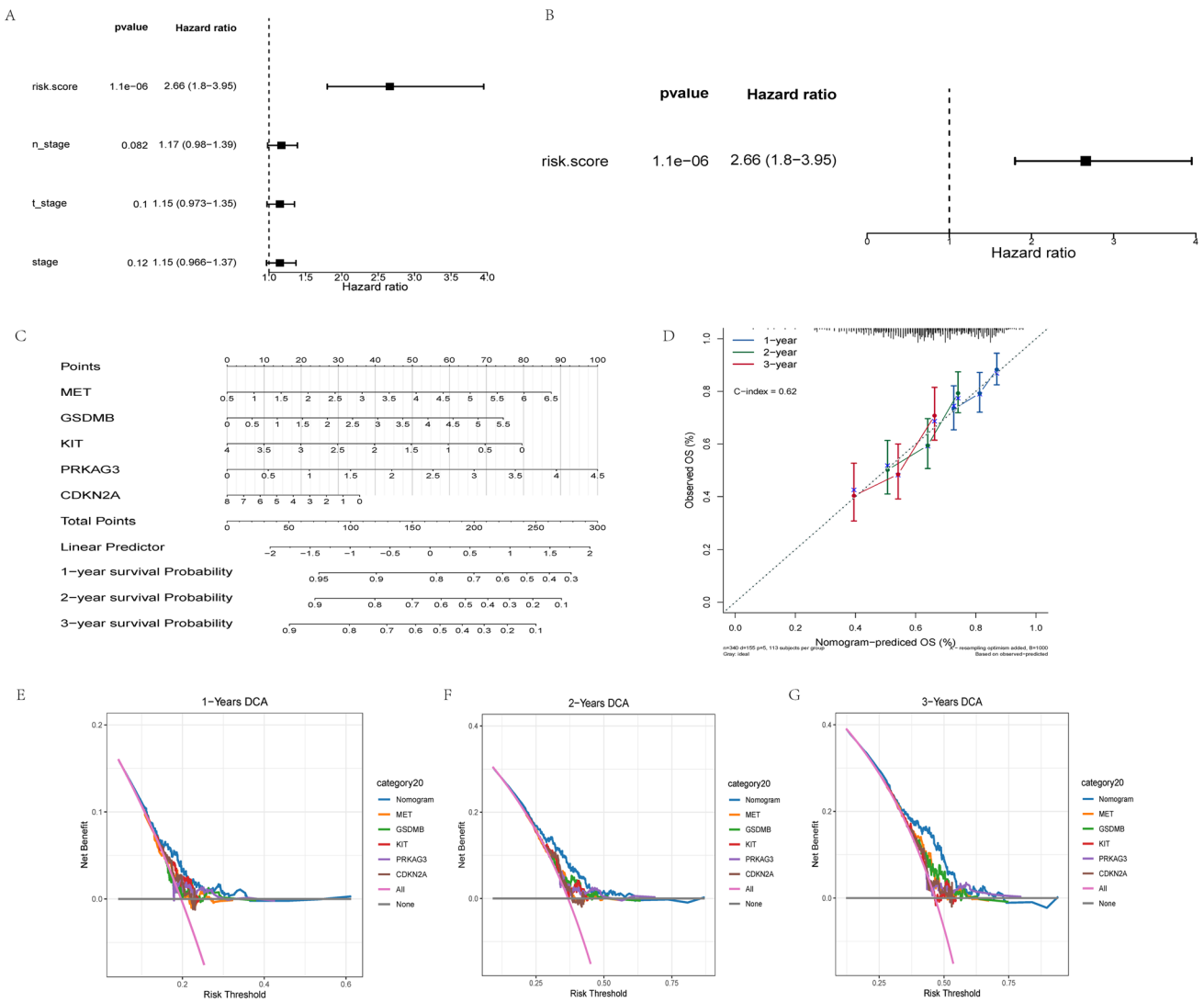


Fig. 4 Development and validation of the prognostic nomogram. **A**, **B** Univariate and multivariate Cox regression analyses of clinical features and risk score. **C** Nomogram for predicting 1-, 2-, and 3-year overall survival. **D** Calibration curves for nomogram predictions. **E**–**G** Decision curve analysis showing clinical benefit of the nomogram at 1-, 2-, and 3-year time points

potentially contributing to OSCC progression. Research shows that GSDMB gene is a part of a prognostic gene signature in OSCC and is associated with cell differentiation and therapeutic response [24, 25]. The gene, when expressed, may influence the programmatic cell death pathways, including pyroptosis, which is a type of programmed cell death distinct from apoptosis. It has been noted to potentially serve as a marker for favorable prognosis in response to treatment [26]. Literature directly associating KIT gene with OSCC in current studies is limited. However, the KIT gene typically plays a role in cell surface receptors in hematopoiesis, melanogenesis, and gametogenesis. In cancer, aberrations in this gene are associated with tumor cell proliferation and survival. Research in similar contexts shows potential impacts in cancer prognosis [24]. Direct associations with OSCC are lacking in recent studies, though it is known that the gene encoding PRKAG3 relates to the energy metabolism pathways, which could influence cancer cell metabolism and survival under stress conditions typically seen in tumors like OSCC [27]. CDKN2A gene is often associated with pathways regulating cell cycle progression, notably influencing the cell's ability to undergo apoptosis in the presence of DNA damage. High expression levels have been linked to better prognosis in head and neck squamous cell carcinomas, suggesting a similar pattern might extend to OSCC [28]. Conversely, under certain conditions, CDKN2A might correlate with poorer outcomes due to its complex role in tumor suppression and survival [29]. These findings underscore the nuanced roles these genes play in OSCC, not only serving as potential biomarkers for prognostic evaluations but also as therapeutic targets to modulate the tumor microenvironment and improve patient outcomes. Each gene's involvement and interaction with

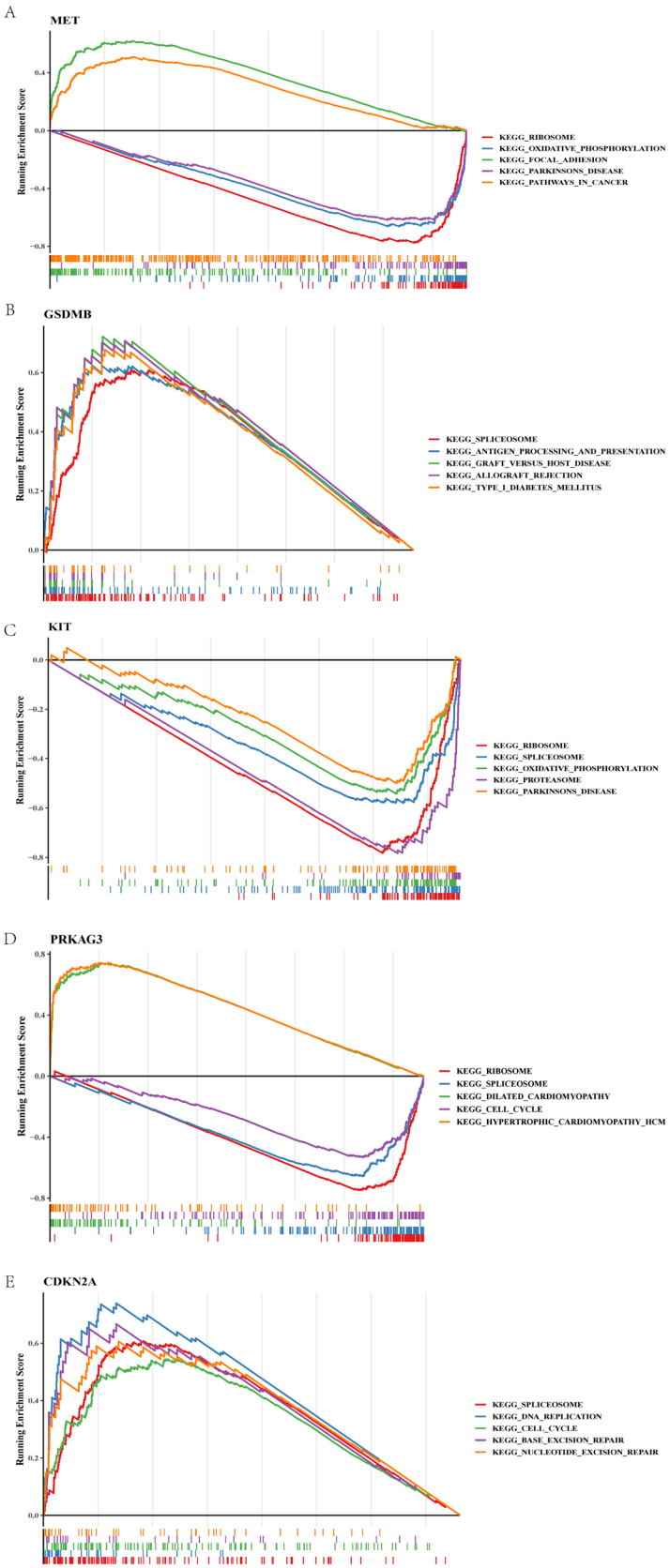


Fig. 5 GSEA pathway analysis of prognostic genes. **A**–**E** Gene Set Enrichment Analysis (GSEA) showing significantly enriched pathways for each prognostic gene: **A** MET, **B** GSDMB, **C** KIT, **D** PRKAG3, and **E** CDKN2 A. Running enrichment scores and gene set member positions are shown for each pathway

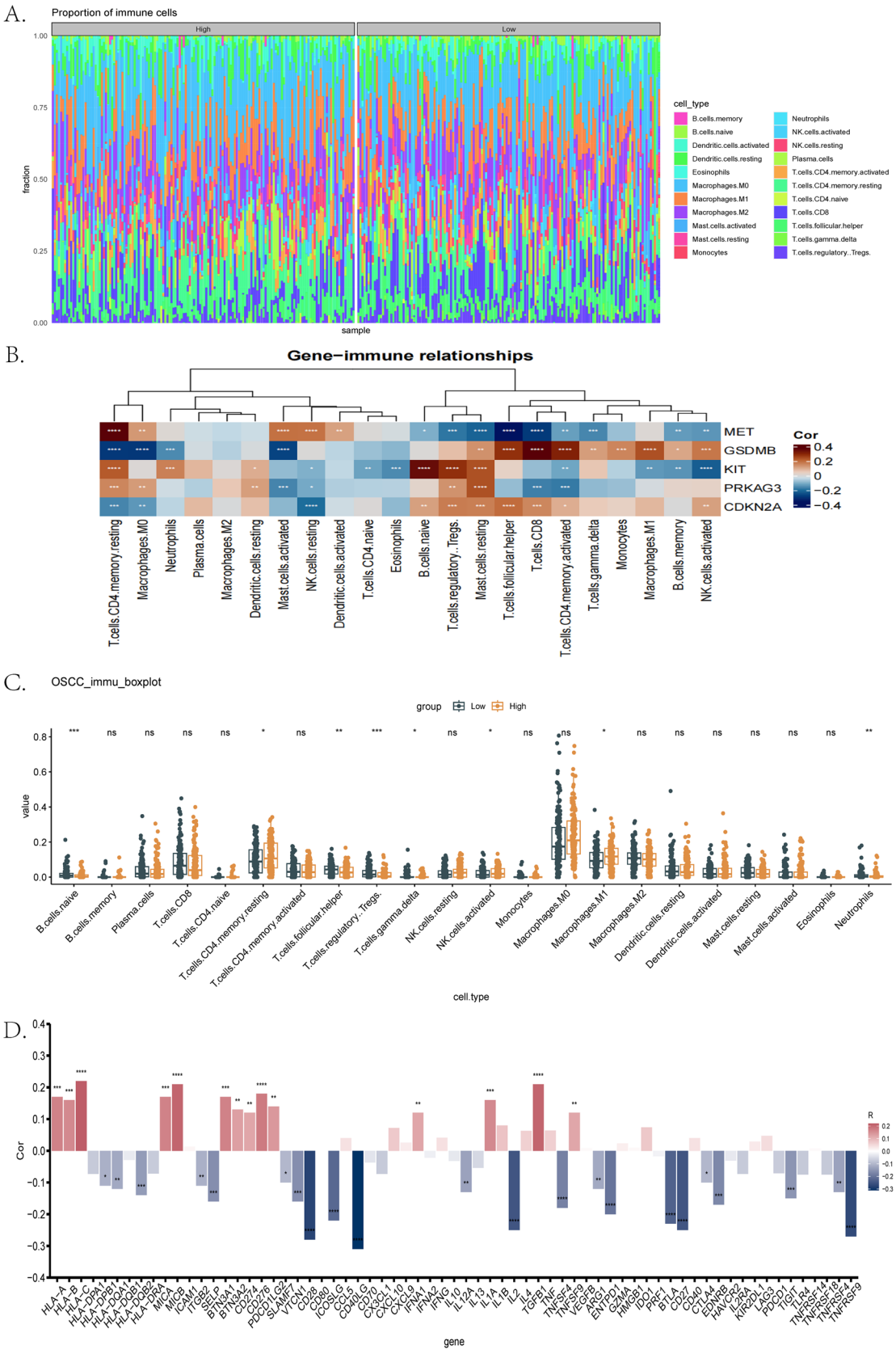


Fig. 6 Analysis of immune cell infiltration in OSCC. **A** Box plots showing the distribution of 22 immune cell types between high- and low-risk groups. **B** Stacked bar plot showing the proportion of immune cells in each sample. **C** Heatmap showing correlations between prognostic genes and immune cell types. **D** Bar chart showing correlations between risk scores and immune-regulatory genes. Asterisks indicate statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

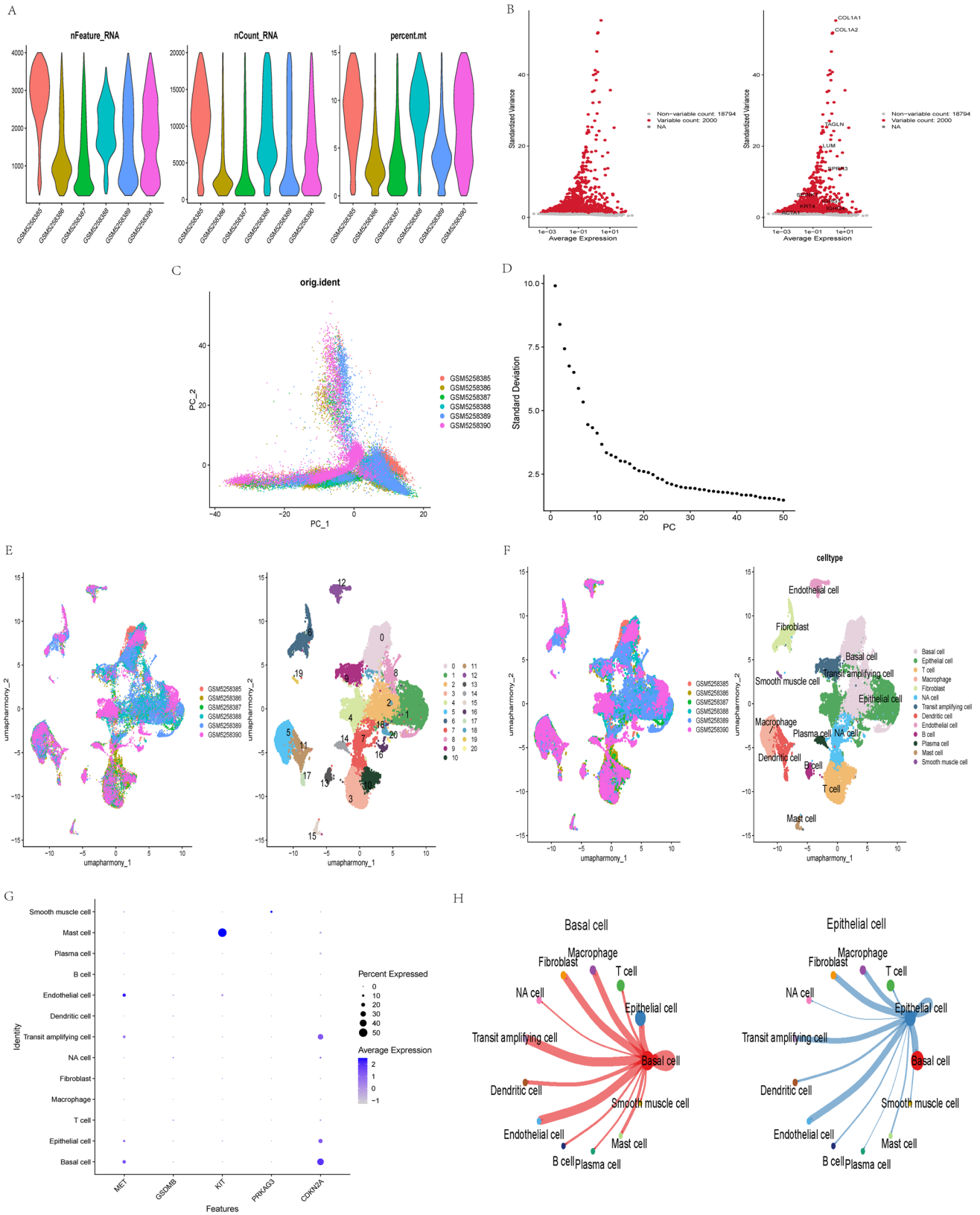


Fig. 7 Single-cell RNA sequencing analysis of OSCC samples. **A** Quality control metrics showing nFeature_RNA, nCount_RNA, and percent.mt distributions. **B** Variable gene analysis plot. **C** PCA plot colored by sample origin. **D** PCA elbow plot. **E**–**F** UMAP visualization colored by sample origin and unsupervised clustering. **G** Expression patterns of prognostic genes across cell types. **H** Cell–cell communication network analysis showing interaction patterns between different cell types

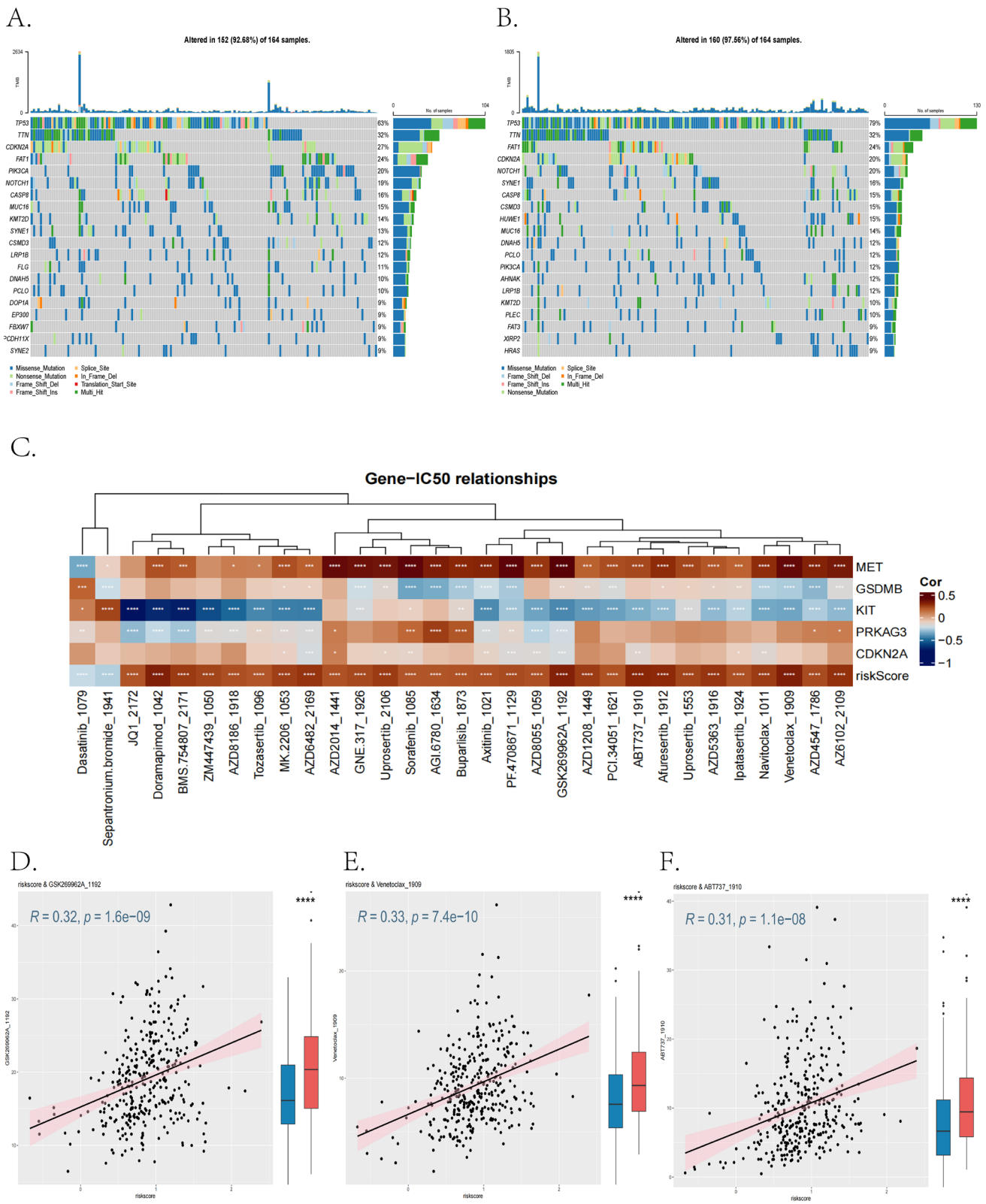


Fig. 8 Somatic mutation profiles and drug sensitivity analysis in OSCC. **A** Oncoplot showing the mutation landscape in the low-risk group. **B** Oncoplot showing the mutation landscape in the high-risk group. **C** Heatmap showing correlations between risk scores and IC50 values of multiple drugs, with color intensity indicating correlation strength. **D**–**F** Correlation analyses between risk scores and sensitivity to three key drugs: **D** GSK2699G2 A_1192, **E** Venetoclax_1909, and **F** ABT737_1910. Each panel includes both a scatter plot showing the relationship between risk score and IC50 values (with high-risk patients in red and low-risk patients in blue) and a box plot comparing IC50 distributions between risk groups. Asterisks indicate statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

programmed cell death pathways provide critical insights into the pathways driving OSCC progression and potential intervention points.

The five genes identified in our prognostic model demonstrate significant correlations with established prognostic factors in OSCC, providing biological context for their prognostic value. MET, known for its involvement in tumor growth and metastasis, has been shown to outperform traditional measures like tumor depth of invasion in predicting lymph node metastasis in early OSCC [30], suggesting its potential utility in refining TNM staging. MET alterations are more frequently observed in HPV-negative tumors [31], potentially serving as an independent prognostic marker that complements HPV status in risk stratification [32, 33]. The GSDMB gene is frequently overexpressed in various tumors and can influence cancer cell behavior. Its role in OSCC, particularly in relation to HPV status and TNM staging, remains to be fully elucidated [34]. In breast cancer models, GSDMB expression did not significantly affect tumor progression, suggesting that its role may be context-dependent and influenced by other factors such as HPV status and tumor stage [34]. KIT, a receptor tyrosine kinase gene, has been associated with distant metastasis and poorer survival rates in OSCC, with its amplification serving as a significant prognostic factor that can provide more accurate prognostic information than TNM staging alone [35]. In the context of HPV status, c-KIT expression in HPV-positive OSCCs can be influenced by tyrosine kinase inhibitors and viral proteins such as E5, which may affect receptor turnover and activation pathways, potentially explaining the distinct therapeutic responses observed in these patients [36, 37]. Furthermore, KIT's role in cancer stemness and epithelial-mesenchymal transition suggests its involvement in tumor growth and metastasis, making it a promising target for personalized treatment strategies, particularly in specific molecular subtypes of OSCC [38, 39]. PRKAG3, which encodes a subunit of AMP-activated protein kinase involved in cellular energy metabolism, has emerged as a significant prognostic factor in our model despite limited direct evidence in previous OSCC studies; the integration of this gene with clinicopathologic factors aligns with approaches shown to improve prognostic accuracy in oral cancer [40]. While not previously identified in HPV-related molecular pathways, PRKAG3 may interact with known altered metabolic processes in OSCC progression, potentially explaining its prognostic value when combined with traditional clinical parameters like TNM staging. CDKN2A demonstrates perhaps the strongest correlation with established prognostic factors, as CDKN2 A copy number loss is associated with poor survival outcomes particularly in advanced TNM stages [41], and its relationship with HPV is well-documented—while HPV-positive OSCC generally shows high p16 (encoded by CDKN2 A) expression and better prognosis, CDKN2 A alterations can still negatively impact outcomes in both HPV-positive and HPV-negative cases [42–44]. These correlations not only validate our prognostic signature but also suggest potential mechanisms through which these genes influence OSCC progression and response to treatment.

As shown in Fig. 5, our GSEA analysis revealed that the five prognostic genes are significantly enriched in distinct yet interconnected pathways that collectively influence OSCC development and progression. Specifically, MET demonstrated significant enrichment in focal adhesion signaling pathways (Fig. 5A), GSDMB was notably enriched in spliceosome pathways (Fig. 5B), KIT showed involvement in oxidative phosphorylation pathways (Fig. 5C), PRKAG3 was enriched in ribosome pathways (Fig. 5D), and CDKN2A exhibited significant enrichment in cell cycle pathways (Fig. 5E). MET gene plays a crucial role in OSCC through its encoded receptor tyrosine kinase, particularly in the Focal Adhesion signaling pathway. Upon binding to its ligand, hepatocyte growth factor (HGF), MET activates multiple downstream signaling pathways, including focal adhesion kinase (FAK), Src kinases, and Rho GTPases, which are essential molecules in cell migration, invasion, and proliferation. Through these signaling pathways, MET enhances the adhesion, migration, and invasion capabilities of OSCC cells, augmenting tumor malignancy and metastatic potential. Consequently, MET serves not only as a significant molecular marker for OSCC but also as a potential target for directed therapy [45, 46]. Similarly, KIT also showed involvement in the oxidative phosphorylation pathway, where it influences mitochondrial function and energy metabolism critical for tumor growth [47]. KIT may regulate mitochondrial DNA expression and oxidative phosphorylation (OXPHOS) activity, with high expression of OXPHOS genes linked to improved survival in head and neck cancers [48]. PRKAG3, which encodes a subunit of AMP-activated protein kinase, was notably enriched in pathways related to ribosome function, highlighting the intricate link between metabolic regulation and protein synthesis in cancer cells [49]. As part of the AMPK complex, PRKAG3 can modulate the mTOR pathway, a key regulator of ribosome biogenesis and protein synthesis [49], thereby influencing the increased protein production demands of rapidly proliferating OSCC cells. GSDMB showed strong enrichment in immune-related pathways and is notably involved in the spliceosome pathway through alternative splicing mechanisms. Different GSDMB isoforms resulting from alternative splicing have distinct functions in pyroptosis and immune evasion [50], with isoforms 3 and 4 capable of inducing pyroptosis when cleaved by Granzyme A from killer lymphocytes, while isoforms 1 and 2 may contribute to tumor survival by evading immune cell-mediated death [51]. CDKN2A exhibited significant enrichment in DNA replication and cell cycle pathways, where it acts as a crucial regulator of G1 to S phase transition. Alterations in CDKN2A disrupt interactions with cyclins and

cyclin-dependent kinases, leading to dysregulated cell proliferation [52]. Moreover, CDKN2A interacts with immune components, influencing immune cell infiltration in tumors [53], which may explain its correlation with immune cell profiles observed in our study.

Several limitations of this study should be acknowledged. First, although our prognostic model was validated in an independent cohort, both the training and validation datasets were retrospectively collected, which may introduce potential selection bias. The validation cohort size of 76 patients was relatively small, and a larger prospective cohort would be necessary to further confirm the robustness of our findings. Second, our study lacked experimental validation of the functional roles of the five identified genes in OSCC progression. In vitro and in vivo experiments would be valuable to elucidate the underlying molecular mechanisms of how these genes influence tumor development and progression. Third, while single-cell RNA sequencing provided valuable insights into cellular heterogeneity, the analysis was limited to only six OSCC samples, which may not fully capture the complexity and diversity of the tumor microenvironment across different patients and disease stages. Furthermore, the clinical information available in our datasets was incomplete for some patients, particularly regarding treatment history and specific therapeutic responses, which limited our ability to assess the model's predictive value for treatment outcomes. Additionally, our study focused solely on transcriptomic data, and the integration of other molecular data types such as genomic, epigenomic, and proteomic information might provide more comprehensive insights. Finally, while our prognostic model showed promising results in predicting survival, its performance in predicting other important clinical outcomes, such as treatment response, recurrence, and metastasis, remains to be evaluated. Future studies addressing these limitations through prospective clinical trials, multi-omics approaches, and comprehensive functional experiments would further strengthen the clinical utility of our findings.

This study has significant implications for both clinical practice and future research directions in OSCC treatment. The five-gene prognostic signature, combined with the nomogram, provides a practical tool for risk stratification and personalized survival prediction in clinical settings. The distinct immune infiltration patterns between risk groups suggest potential applications in guiding immunotherapy decisions, particularly in patient selection for immune checkpoint inhibitor treatment. The identification of strong stromal-epithelial cell interactions through single-cell analysis reveals potential therapeutic targets within the tumor microenvironment. Future research should focus on prospective clinical trials to validate this prognostic model's utility, particularly its ability to predict treatment responses. Investigation of the molecular mechanisms underlying these five genes' roles in OSCC progression could reveal new therapeutic targets. Integration of this signature with other clinical parameters and molecular markers might lead to more comprehensive prognostic models. The patterns observed in cell–cell communication warrant further investigation into targeting specific cellular interactions for therapeutic purposes.

5 Conclusion

In this study, we established and validated a novel five-gene prognostic signature based on programmed cell death-related genes (MET, GSDMB, KIT, PRKAG3, and CDKN2A) that effectively predicts survival in OSCC patients. Integration of single-cell RNA sequencing analysis revealed the predominant expression of these prognostic genes in stromal and epithelial cells, highlighting the importance of tumor-stromal interactions in disease progression. The prognostic model demonstrated robust performance in both training and validation cohorts, and the constructed nomogram showed excellent predictive accuracy for individual survival probability. Drug sensitivity analysis identified potential therapeutic compounds associated with the risk score, providing new insights for personalized treatment strategies. Together, these findings not only advance our understanding of OSCC progression but also provide a promising tool for risk stratification and treatment optimization in clinical practice.

Author contributions YL conceived and designed the study, performed the bioinformatics analyses, interpreted the results, and drafted the manuscript. YL also supervised the entire project and was responsible for the final approval of the manuscript as the corresponding author. YY contributed to data collection, assisted with statistical analyses, and helped with the interpretation of results. SH participated in data processing, visualization of results, and revision of the manuscript. SL contributed to data collection, assisted with statistical analyses, contributed to the revision of the manuscript, and served as the senior author. All authors read and approved the final manuscript.

Funding The study was self-funded by the authors.

Data availability All datasets utilized in the study are publicly available.

Code availability The codes during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval Not applicable.

Competing interests The authors declare no competing interests.

References

1. Wei S, Han C, Mo S, Huang H, Luo X. Advancements in programmed cell death research in antitumor therapy: a comprehensive overview. Apoptosis. 2025;30(1):401–21.
2. Shao S-H, Qiu S-T, Li Y, Zhang Y-Y, Cao Y-J, DI C-H, et al. Ferroptosis: mechanisms and application in tumor treatment. Zhongguo yi xue ke xue yuan xue bao Acta Academiae Medicinae Sinicae. 2024;46(5):732–9.
3. Fei C, Zhen X, Shiqiang Z, Jun P. Frontier knowledge and future directions of programmed cell death in clear cell renal cell carcinoma. Cell Death Discov. 2024;10(1):113.
4. Newton K, Strasser A, Kayagaki N, Dixit VM. Cell death. Cell. 2024;187(2):235–56.
5. Liu J, Hong M, Li Y, Chen D, Wu Y, Hu Y. Programmed cell death tunes tumor immunity. Front Immunol. 2022;13:847345.
6. Peng F, Liao M, Qin R, Zhu S, Peng C, Fu L, et al. Regulated cell death (RCD) in cancer: key pathways and targeted therapies. Signal Transduct Target Ther. 2022;7(1):286.
7. Yu L, Huang K, Liao Y, Wang L, Sethi G, Ma Z. Targeting novel regulated cell death: ferroptosis, pyroptosis and necroptosis in anti-PD-1/PD-L1 cancer immunotherapy. Cell Prolif. 2024;57(8):e13644.
8. Shkarina K, Broz P, editors. Selective induction of programmed cell death using synthetic biology tools. Seminars Cell Dev Biol. 2024;156:74–92.
9. Deng Y, Li K, Yan W, Li K, Wang C. Gene expression regulation and the signal transduction of programmed cell death. Curr Issues Mol Biol. 2024;46(9):10264.
10. Wang M, Yu F, Zhang Y, Li P. Programmed cell death in tumor immunity: mechanistic insights and clinical implications. Front Immunol. 2024;14:1309635.
11. Zhou L, Wu Y, Ying Y, Ding Y. Current knowledge of ferroptosis in the pathogenesis and prognosis of oral squamous cell carcinoma. Cellular Signal. 2024;119:111176.
12. Ai Y, Meng Y, Yan B, Zhou Q, Wang X. The biochemical pathways of apoptotic, necroptotic, pyroptotic, and ferroptotic cell death. Mol Cell. 2024;84(1):170–9.
13. Ning J, Chen L, Zeng Y, Xiao G, Tian W, Wu Q, et al. The scheme, and regulative mechanism of pyroptosis, ferroptosis, and necroptosis in radiation injury. Int J Biol Sci. 2024;20(5):1871.
14. Zhang Q, Fan X, Zhang X, Ju S. Ferroptosis in tumors and its relationship to other programmed cell death: role of non-coding RNAs. J Transl Med. 2023;21(1):514.
15. Kong Y, Li J, Lin R, Lu S, Rong L, Xue Y, et al. Understanding the unique mechanism of ferroptosis: a promising therapeutic target. Front Cell Dev Biol. 2024;11:1329147.
16. Fernández-Lázaro D, Sanz B, Seco-Calvo J. Mechanisms of programmed cell death: structural and functional pathways. A narrative review. Invest Clin. 2024;65(2):230–52.
17. Pang Q, Tang Z, Luo L. The crosstalk between oncogenic signaling and ferroptosis in cancer. Crit Rev Oncol Hematol. 2024;197:104349.
18. Tomczak K, Czerwińska P, Wiznerowicz M. Review the cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol. 2015;2015(1):68–77.
19. Clough E, Barrett T. The gene expression omnibus database. Stat Genom Methods Protocols. 2016;93–110.
20. Harrell FE Jr, Harrell MFE Jr, Hmisc D. Package 'rms.' Vanderbilt Univ. 2017;229(201):7.
21. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. JAMA. 2015;313(4):409–10.
22. Verma M. Unravelling immune cell signatures: CIBERSORTx-assisted construction of signature matrix from single-cell data. bioRxiv. 2024:2024.05. 05.592045.
23. Wang J, Ouyang H, Zhang C, Li S, Xiang J. A novel intelligent global harmony search algorithm based on improved search stability strategy. Sci Rep. 2023;13(1):7705.
24. Xin Y, Zhang J, Jiang Q, Qiu J. Construction of prognostic signature of patients with oral squamous cell carcinoma based on pyroptosis-related long non-coding RNAs. Front Surg. 2022;9:935765.
25. Wang A, Zhang C, Wang Y, Diao P, Cheng J. Leveraging programmed cell death patterns to predict prognosis and therapeutic sensitivity in OSCC. Oral Diseases.

26. Wang S, Zhang M, Wu Z, Zhu S, Wan S, Zhang B, et al. GSDME is related to prognosis and response to chemotherapy in oral cancer. J Dent Res. 2022;101(7):848–58.
27. Wang F, Fang L, Fu B, Fan C. Construction of a prognostic risk assessment model for HER2+ breast cancer based on autophagy-related genes. Breast Cancer. 2023;30(3):478–88.
28. Fan K, Dong Y, Li T, Li Y. Cuproptosis-associated CDKN2A is targeted by plicamycin to regulate the microenvironment in patients with head and neck squamous cell carcinoma. Front Genet. 2023;13:1036408.
29. Zhang D, Wang T, Zhou Y, Zhang X. Comprehensive analyses of cuproptosis-related gene CDKN2A on prognosis and immunologic therapy in human tumors. Medicine. 2023;102(14):e33468.
30. De Herdt MJ, van der Steen B, van der Toom QM, Aaboubout Y, Willems SM, Wieringa MH, et al. The potential of MET immunoreactivity for prediction of lymph node metastasis in early oral tongue squamous cell carcinoma. Front Oncol. 2021;11:638048.
31. Akashi K, Ebihara Y, Omura G, Saito Y, Yoshida M, Ando M, et al. Frequent copy gain of the MET gene in hypopharyngeal and laryngeal cancer in the japanese population. J Cancer Ther. 2015;6(12):1093–102.
32. Lohavanichbutr P, Méndez E, Holsinger FC, Rue TC, Zhang Y, Houck J, et al. A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. Clin Cancer Res. 2013;19(5):1197–203.
33. Liu J, Jackson C, Anees A, Noor Z, Williams S, Porceddu S, et al. A proteomic signature associated with prognosis in HPV-related locally advanced oropharyngeal squamous cell carcinoma (LA-OPSCC). J Clin Oncol. 2023;41:6055.
34. Sarrio D, Rojo-Sebastián A, Teijo A, Pérez-López M, Díaz-Martín E, Martínez L, et al. In vivo analysis of the role of gasdermin-b (GSDMB) in cancer using novel knock-in mouse models. bioRxiv. 2021:2021.05. 27.445936.
35. Oikawa Y, Morita Ki, Kayamori K, Tanimoto K, Sakamoto K, Katoh H, et al. Receptor tyrosine kinase amplification is predictive of distant metastasis in patients with oral squamous cell carcinoma. Cancer Sci. 2017;108(2):256–66.
36. Kramer B, Kneissle M, Birk R, Rotter N, Aderhold C. Tyrosine kinase inhibition in HPV-related squamous cell carcinoma reveals beneficial expression of cKIT and SRC. Anticancer Res. 2018;38(5):2723–31.
37. Cortelazzi B, Verderio P, Ciniselli CM, Pizzamiglio S, Bossi P, Gloghini A, et al. Receptor tyrosine kinase profiles and human papillomavirus status in oropharyngeal squamous cell carcinoma. J Oral Pathol Med. 2015;44(9):734–45.
38. Sheikh E, Tran T, Vranic S, Levy A, Bonfil RD. Role and significance of c-KIT receptor tyrosine kinase in cancer: a review. Bosn J Basic Med Sci. 2022;22(5):683.
39. Wagner S, Würdemann N, Hübbers C, Reuschenbach M, Prigge E-S, Wichmann G, et al. HPV-assoziierte Kopf-Hals-Karzinome: Mutationssignaturen und genomische Aberrationen (Leitthema). HNO. 2015;63(11):758–67.
40. Chang S-W, Kareem SA, Merican AFMA, Zain RB. A hybrid prognostic model for oral cancer based on clinicopathologic and genomic markers. Sains Malaysiana. 2014;43(4):567–73.
41. Chen WS, Bindra RS, Mo A, Hayman T, Husain Z, Contessa JN, et al. CDKN2A copy number loss is an independent prognostic factor in HPV-negative head and neck squamous cell carcinoma. Front Oncol. 2018;8:95.
42. Lee J, Hong D, Lee EK, Jung Y-S, Park WS. Recurrent mutations of CDKN2A genes characterize worse prognosis of HPV positive and negative oropharyngeal squamous cell carcinoma. Cancer Res. 2016;76(14_Supplement):111.
43. Shponka IS I, Bondarenko O, Kovtunenko O, Rakhmanov V. Deletion of cyclin dependent kinase inhibitor 2a gene as a marker of oropharyngeal carcinomas nonassociated with human papillomavirus and its prognostic value. Медичні перспективи= Medicni perspektivi (Medical perspectives). 2024;29(2):56–61.
44. Wang Y, Zhou C, Li T, Luo J. Prognostic value of CDKN2A in head and neck squamous cell carcinoma via pathomics and machine learning. J Cell Mol Med. 2024;28(9):e18394.
45. Yokokawa M, Morita Ki, Oikawa Y, Kayamori K, Sakamoto K, Ikeda T, et al. Co-expression of EGFR and MET has a synergistic effect on the prognosis of patients with oral squamous cell carcinoma. J Oral Pathol Med. 2020;49(3):235–42.
46. Freudlsperger C, Alexander D, Reinert S, Hoffmann J. Prognostic value of c-Met expression in oral squamous cell carcinoma. Experimental Therapeutic Med. 2010;1(1):69–72.
47. Frederick M, Skinner HD, Kazi SA, Sikora AG, Sandulache VC. High expression of oxidative phosphorylation genes predicts improved survival in squamous cell carcinomas of the head and neck and lung. Sci Rep. 2020;10(1):6380.
48. Chattopadhyay E, De Sarkar N, Singh R, Ray A, Roy R, Paul RR, et al. Genome-wide mitochondrial DNA sequence variations and lower expression of OXPHOS genes predict mitochondrial dysfunction in oral cancer tissue. Tumor Biol. 2016;37:11861–71.
49. Kusnadi E, Trigos A, Cullinane C, Goode D, Larsson O, Devlin J, et al. Reprogrammed mRNA translation drives metabolic response to therapeutic targeting of ribosome biogenesis. bioRxiv. 2019:847723.
50. Kong Q, Xia S, Pan X, Ye K, Li Z, Li H, et al. Alternative splicing of GSDMB modulates killer lymphocyte–triggered pyroptosis. Sci Immunol. 2023;8(82):eadg3196.
51. Oltra SS, Colomo S, Sin L, Pérez-López M, Lázaro S, Molina-Crespo A, et al. Distinct GSDMB protein isoforms and protease cleavage processes differentially control pyroptotic cell death and mitochondrial damage in cancer cells. Cell Death Differ. 2023;30(5):1366–81.
52. Hsieh R, Nico MM, Coutinho-Camillo CM, Buim ME, Sangueza M, Lourenço SV. The CDKN2A and MAP kinase pathways: molecular roads to primary oral mucosal melanoma. Am J Dermatopathol. 2013;35(2):167–75.
53. Soltan MA, Alhanshani AA, Shati AA, Alqahtani YA, Alshaya DS, Alharthi J, et al. Cyclin dependent kinase inhibitor 2A genetic and epigenetic alterations interfere with several immune components and predict poor clinical outcome. Biomedicines. 2023;11(8):2254.