Supplementary Material

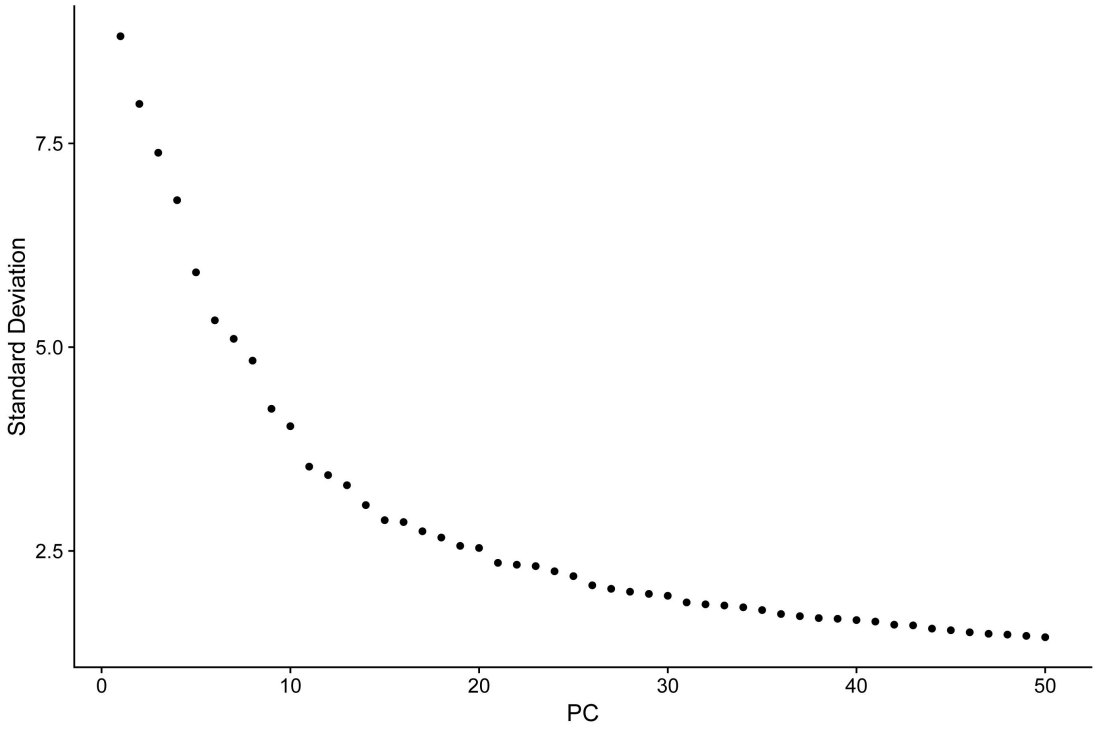

Figure S1 The scree plot illustrates the standard deviation across principal components (PCs) obtained through principal component analysis (PCA). The x-axis represents the principal components, while the y-axis denotes their respective standard deviation values. A decreasing trend in standard deviation is observed as the component index increases, indicating that components with lower indices capture the primary variance within the data. By identifying the point at which the standard deviation shows a marked reduction, an informed selection of the principal components for downstream analysis can be made.

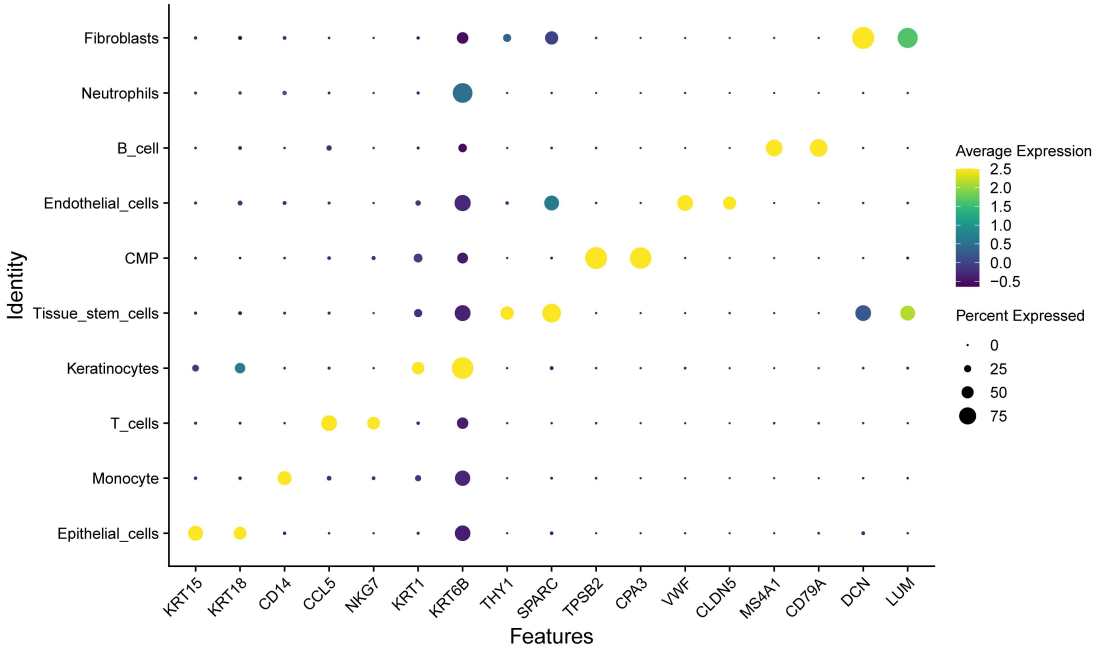

Figure S2 Marker genes verification results of different cell groups

1.1 Pseudotime, Transcription Factor Prediction and InferCNV Analysis Results

In this study, we utilized the DoRothEA package to explore the regulation of different OSCC samples by transcription factors (TFs) (Figure S3A). Enrichment analyses of TFs using GO, KEGG, and Metascape revealed their involvement in pathways relevant to OSCC, such as gland development, hematopoiesis, cell proliferation, and dysregulated transcription in cancer (Figure S3B-D). Additionally, pseudotime analysis was employed to determine the distribution of different cell types at various developmental stages (Figure S3E-G). Keratinocytes were predominantly located in the pre-differentiation stage, while monocytes and tissue stem cells were primarily situated in the post-differentiation stage.

Subsequently, we conducted an inferCNV analysis of cells within oral cancer tissues, using T cells as a reference group to explore the CNV patterns across different cell populations. The results, presented in Figure S4, show that the reference group (T cells) exhibits a stable genomic expression pattern, while the observed cell populations (including epithelial cells, neutrophils, fibroblasts, etc.) display significant copy number variation features. Specifically, we observed copy number gains (in red) or losses (in blue) in multiple genomic regions. Notably, epithelial cells demonstrated a strong CNV signal, whereas immune cells such as neutrophils and monocytes exhibited relatively stable or weaker CNV signals. These findings suggest that there may be significant genomic differences among various cell types within the tumor microenvironment.

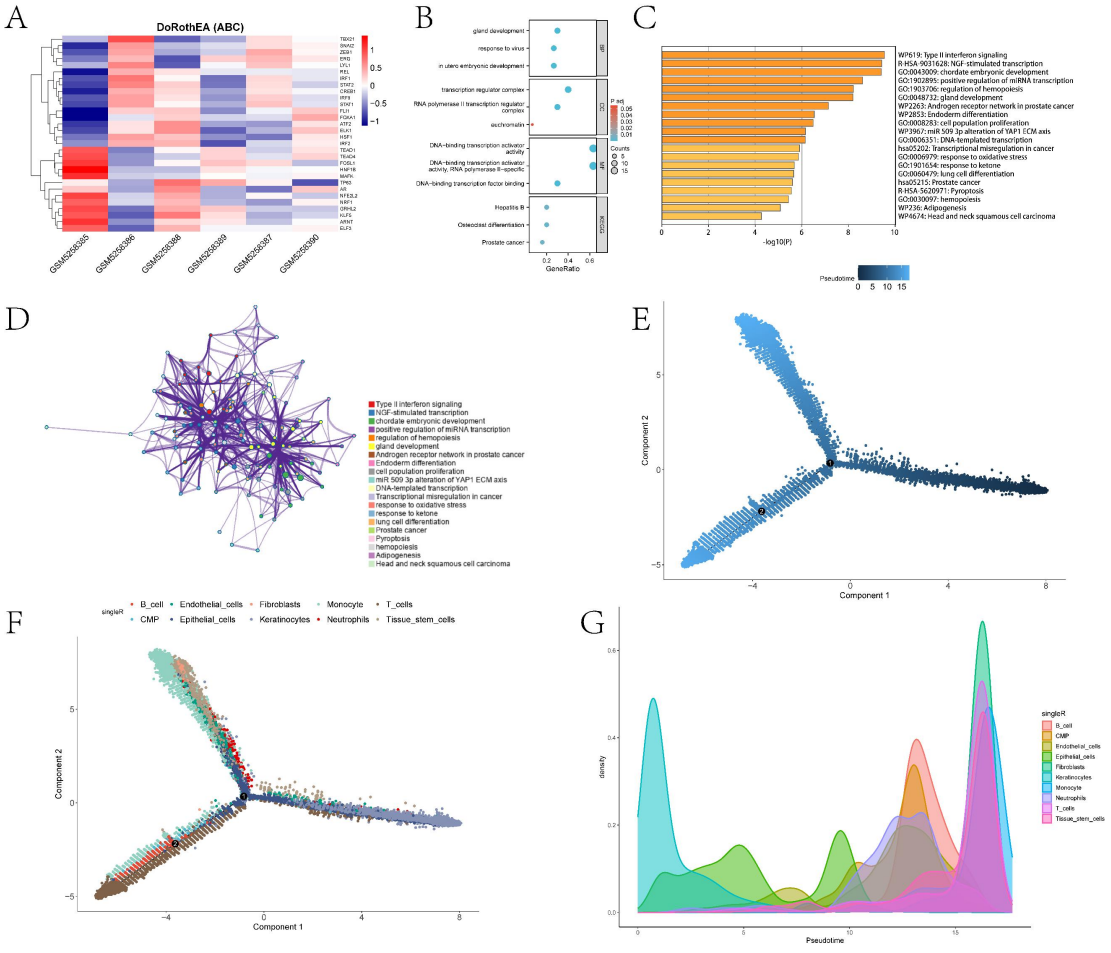

Figure S3 TF and pseudotime analysis results. A depicts the heatmap of TF activity across different samples. B illustrates the results of GO and KEGG enrichment analysis for TFs. C

displays the enrichment analysis results of TFs from Metascape. D shows the pathway interaction network obtained from Metascape enrichment. E and F represent the distribution of cells grouped by pseudotime and type, respectively. G is the density plot of different cell types along the pseudotime axis.

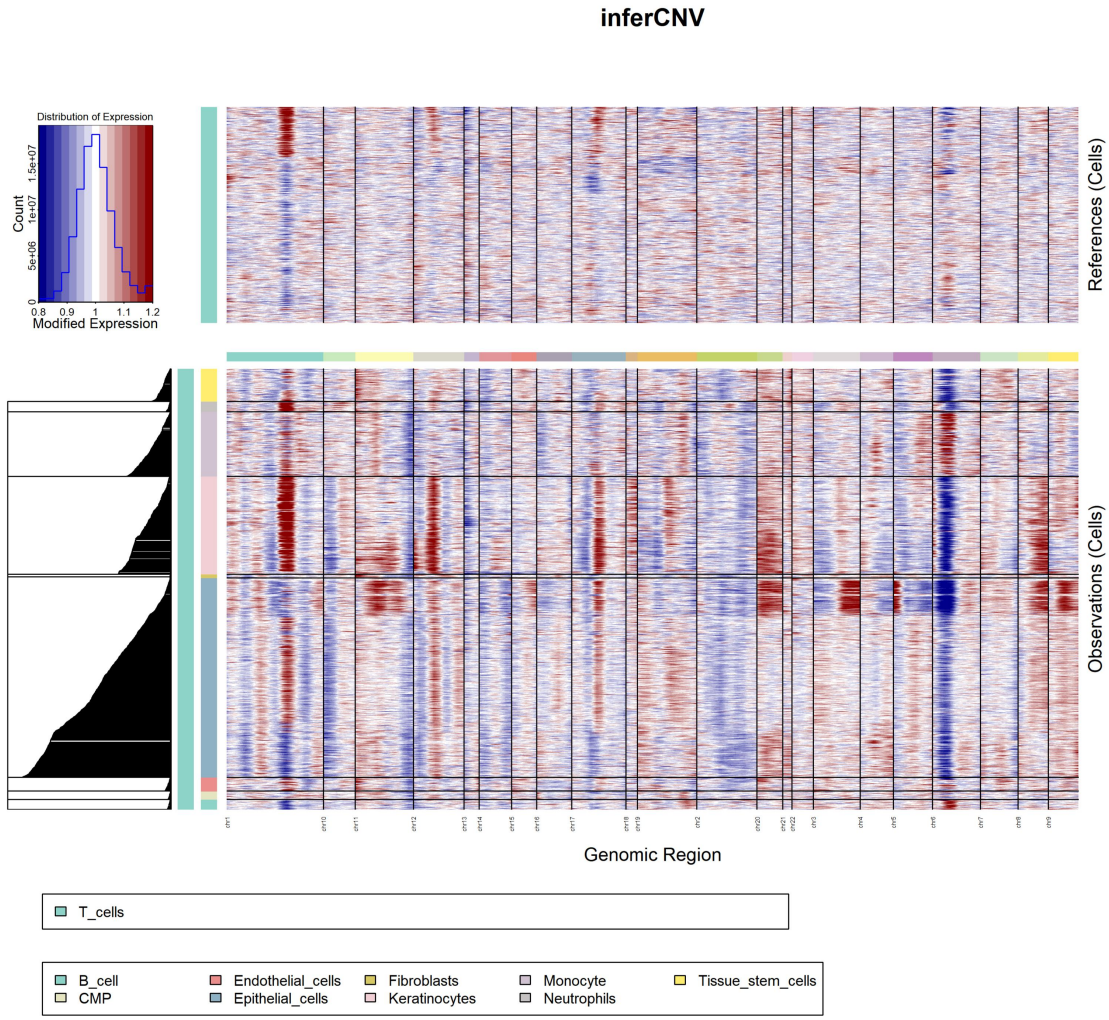

Figure S4 Hierarchical heatmap of large-scale CNVs in different cell types

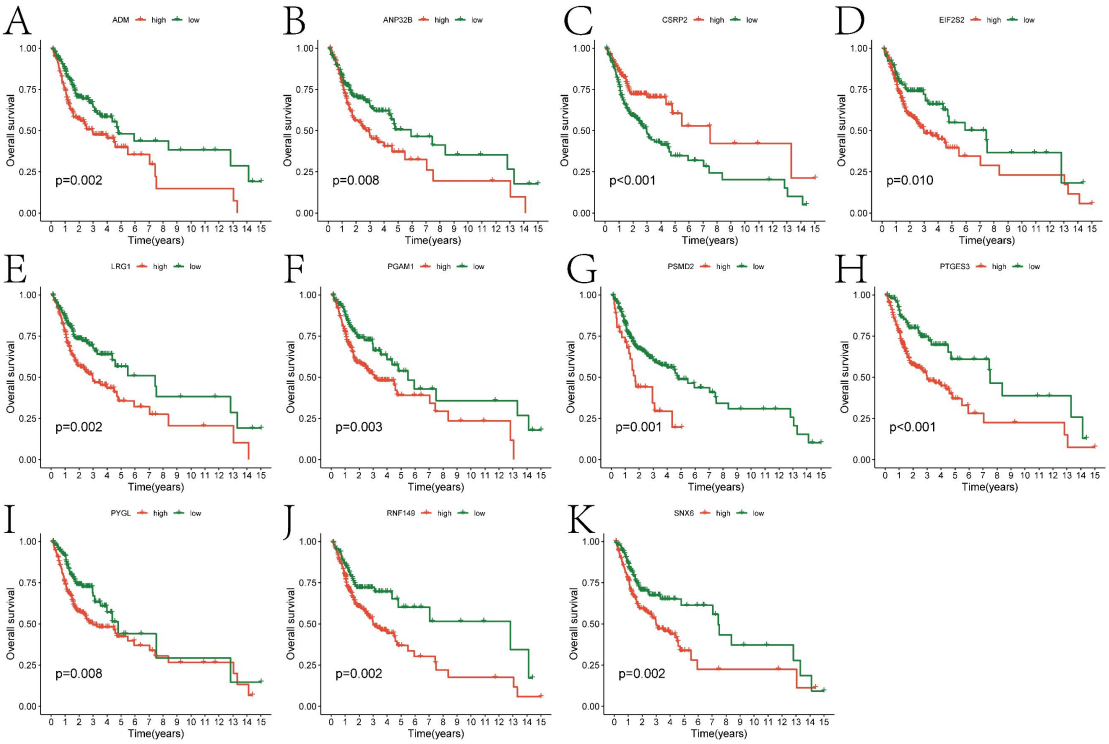

Figure S5 KM survival curve of prognostic related genes. A-K gave the KM survival curves of 11 prognostic genes.

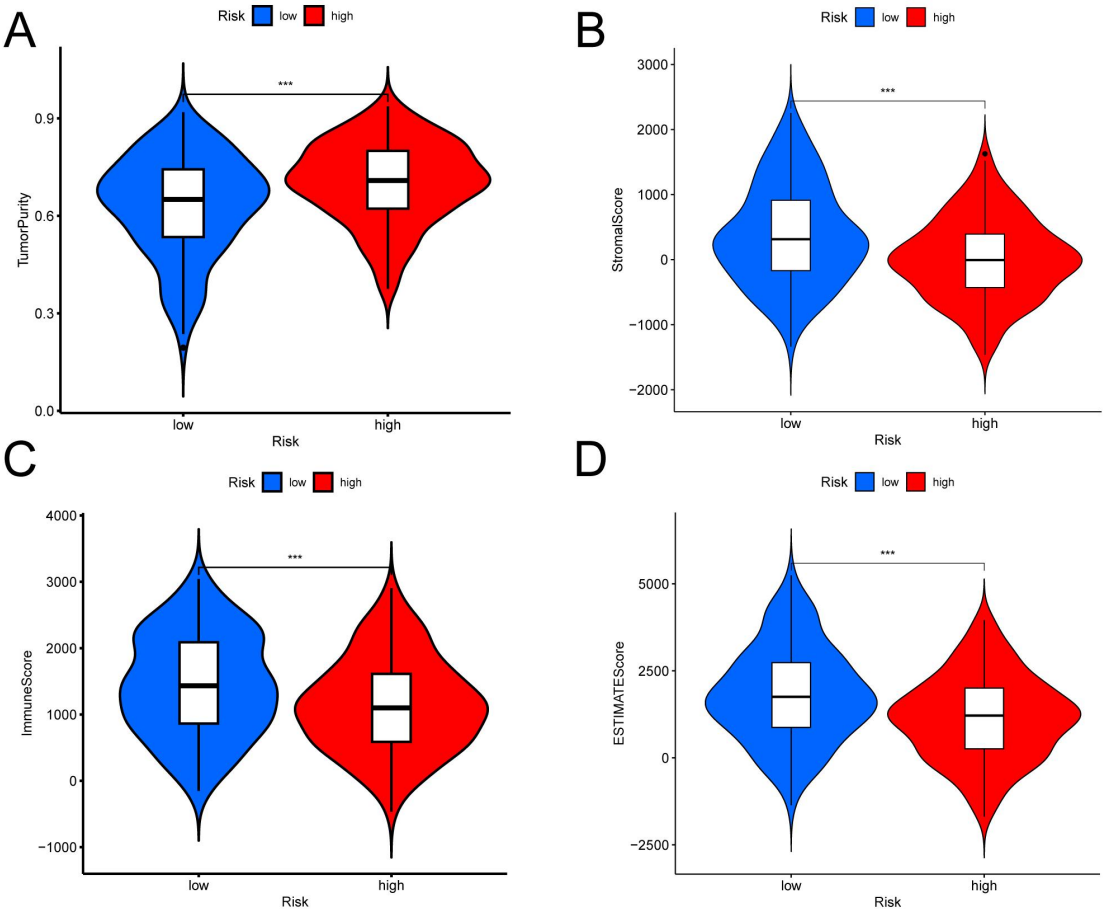

Figure S6 The analysis results of E2F7 and CDH4 expression in neutrophils and their correlation with neutrophil infiltration abundance. Panel A shows a bubble plot of E2F7 and CDH4 expression in neutrophils, while Panel B presents the correlation analysis results between E2F7 and CDH4 and neutrophil infiltration abundance.

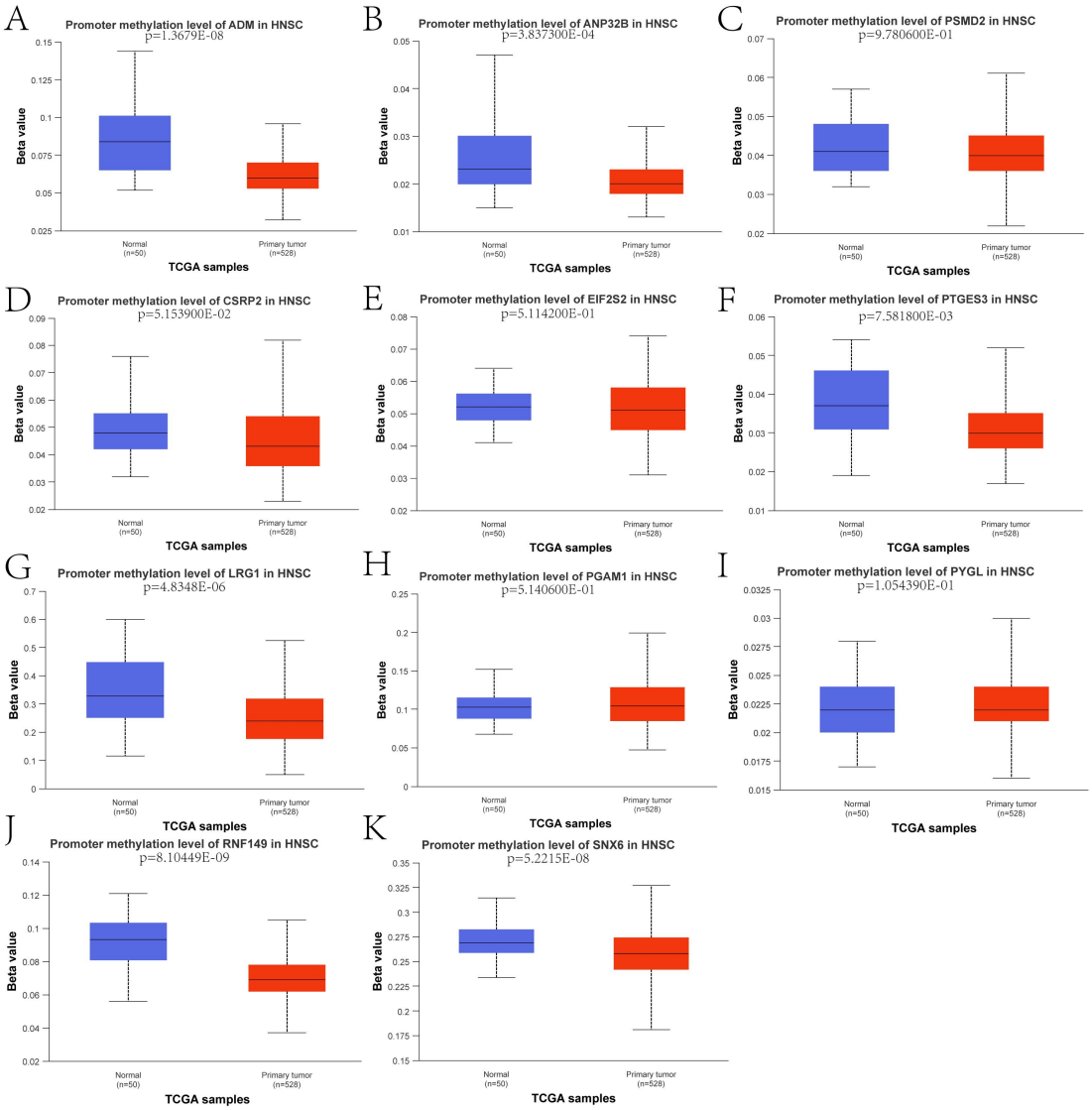

Figure S7 The methylation levels of prognostic genes in the control group and the OSCC group. A-K The methylation levels of ADM, ANP32B, CSRP2, EIF2S2, LRG1, PGAM1, PSMD2, PTGES3, PYGL, RNF149, and SNX6 in both groups.

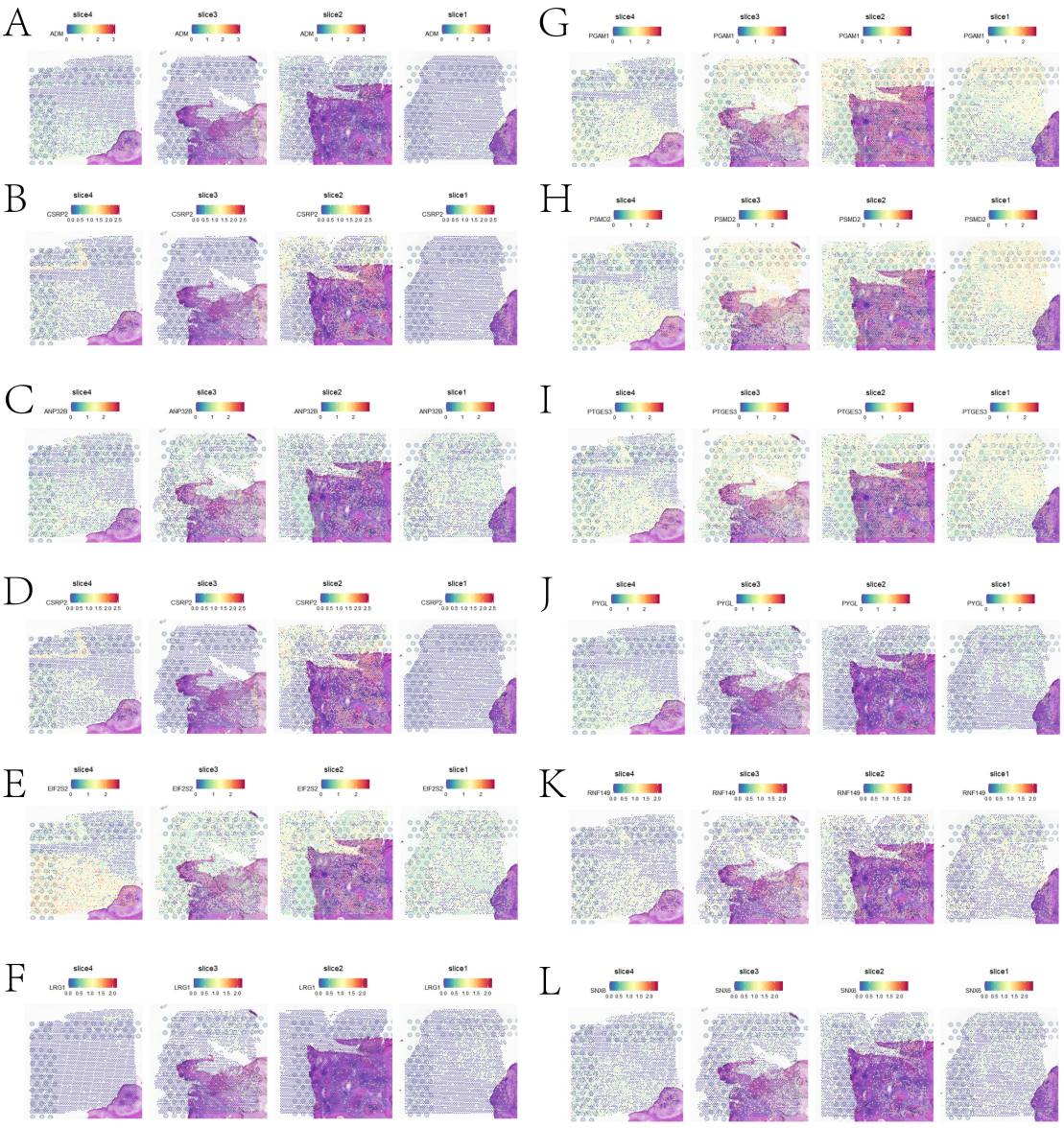

Figure S8 The expression profiles of the prognostic genes in four OSCC tissue sections. Panels A through K represent the heatmaps of ADM, ANP32B, CSRP2, EIF2S2, LRG1, PGAM1, PSMD2, PTGES3, PYGL, RNF149, and SNX6 expression, respectively, across the tissue sections.

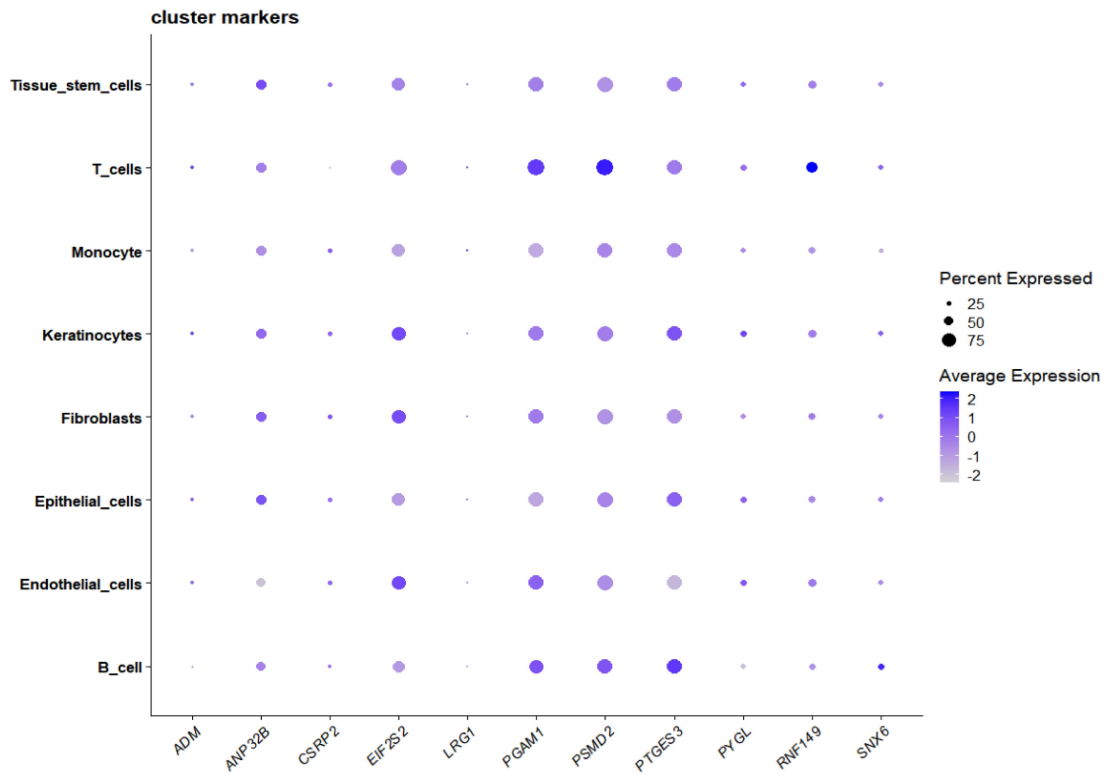

Figure S9 The expression profiles of the prognostic genes in different cell types.

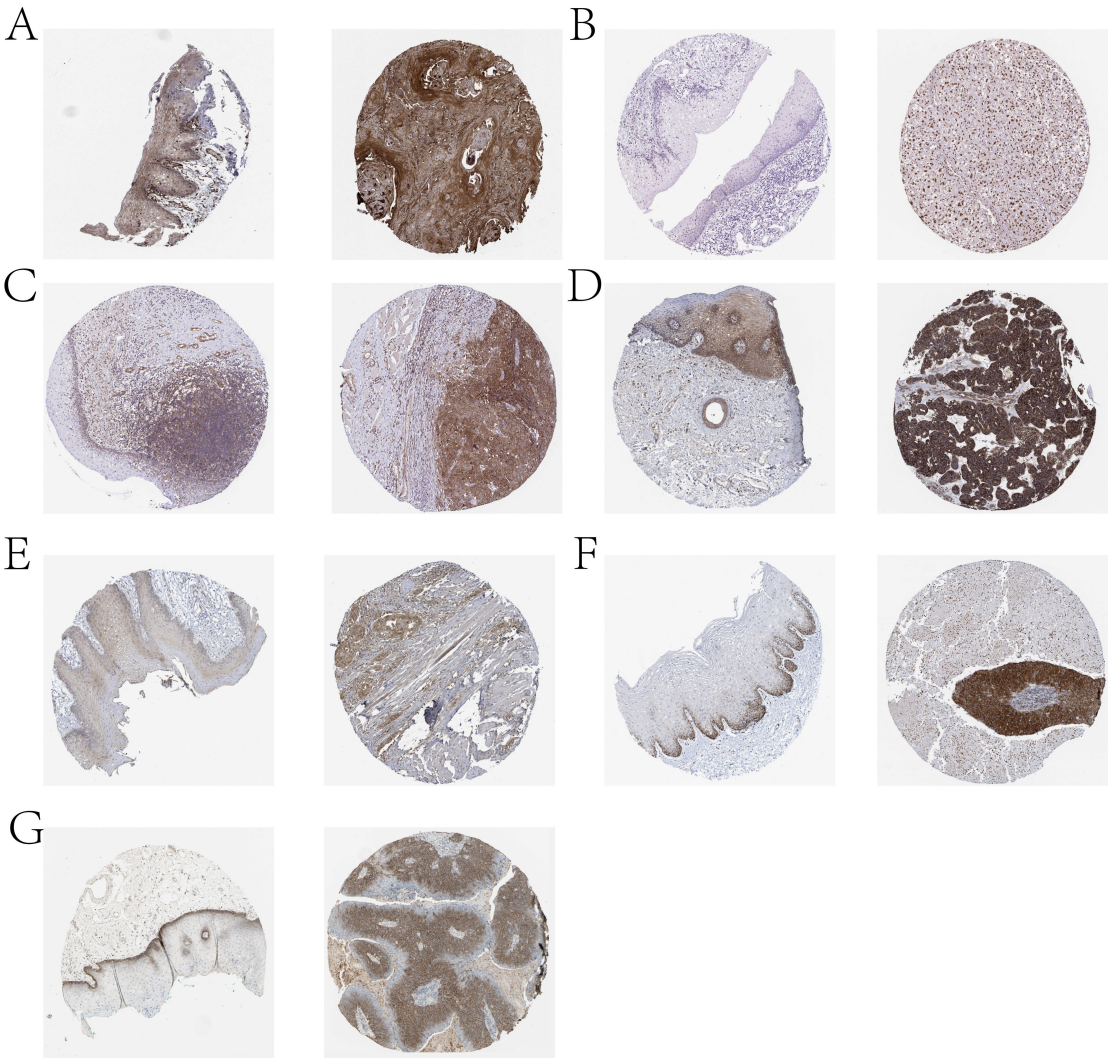

Figure S10 Immunohistochemical results of selected prognostic genes in the control and OSCC groups are as follows: A. Immunohistochemical images of ADM show medium staining in the control group (left) and high staining in the OSCC group (right). B. Images of ANP32B demonstrate low staining in the control group (left) and high staining in the OSCC group (right). C. CSRP2 exhibits low staining in the control group (left) and high staining in the OSCC group (right). D. EIF2S2 has medium staining in the control group (left) and high staining in the OSCC group (right). E. PGAM1 shows low staining in the control group (left) and medium staining in the OSCC group (right). F. PTGES3 demonstrates low staining in the control group (left) and high staining in the OSCC group (right). G. RNF149 exhibits low staining in the control group (left) and medium staining in the OSCC group (right).