Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Integrative single-cell and bulk transcriptome analyses reveal a carnitine-related prognostic signature and fibroblast plasticity in HNSCC

Leyin Hu

m15058440925_3@163.com

Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences

Bing Wang

Shidong Hospital, Yangpu District

Shengrong Dong[1]

Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences

Shimeng Jiang[1]

Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences

Rong Xiong[1]

Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences

---

Research Article

Keywords: Head and neck squamous cell carcinoma, Carnitine, Risk model, Immune infiltration, Single-cell RNA sequencing

Posted Date: October 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11128796/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Integrative single-cell and bulk transcriptome analyses reveal a carnitine-related prognostic signature and fibroblast plasticity in HNSCC

Leyin Hu†*[1], Bing Wang†[2], Shengrong Dong[1], Shimeng Jiang[1], Rong Xiong[1]

[1] Department of Pathology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China

[2]Department of Pathology, Shidong Hospital, Yangpu District, Shanghai 200438, China

† These authors contributed equally to this work.

*Corresponding author. Email: m15058440925_3@163.com

Abstract

Background: Dysregulated carnitine metabolism contributes to tumor metabolic reprogramming, but the prognostic value of carnitine-related genes (CRGs) in head and neck squamous cell carcinoma (HNSCC) remains unclear.

Methods: Differentially expressed genes (DEGs) were identified in The Cancer Genome Atlas (TCGA)-HNSCC cohort and intersected with CRGs to obtain candidate genes. Prognostic genes were screened by univariate and multivariate Cox regression to construct a risk model, evaluated by time-dependent receiver operating characteristic (ROC) curves and Kaplan-Meier (KM) analysis, and externally validated in GSE41613. CIBERSORT was used to compare immune cell infiltration between the high-risk group (HRG) and low-risk group (LRG), and single-cell RNA sequencing (scRNA-seq) data (GSE181919) were analyzed to identify the key cell type associated with prognostic gene expression.

Results: Intersection of 4114 DEGs with 91 CRGs yielded 26 candidate genes. Six prognostic genes (ACADL, SERPINA5, PRKAA2, ACACB, SLC25A20, GPR12) were identified; the resulting risk model achieved 1-, 3-, and 5-year area under the curve (AUC) values of 0.65, 0.66, and 0.65 in the training set and 0.70, 0.64, and 0.60 in external validation, remaining an independent prognostic factor ($P < 0.001$). CD8+ T cells and regulatory T cells were significantly decreased in the HRG ($P < 0.01$). Single-cell analysis identified fibroblasts as the dominant signaling hub; five prognostic genes were predominantly expressed in fibroblasts and progressively downregulated along pseudotime.

Conclusions: A six-gene CRG-based risk model was constructed and externally validated for HNSCC prognosis, showing associations with immune infiltration and fibroblast-mediated microenvironment remodeling, offering a potential reference for prognostic stratification.

Keywords: Head and neck squamous cell carcinoma; Carnitine; Risk model; Immune infiltration; Single-cell RNA sequencing

1. Introduction

Head and neck squamous cell carcinoma (HNSCC), a malignant epithelial neoplasm arising from the mucosal lining of the oral cavity, pharynx, and larynx, represents the sixth most common malignancy worldwide, with over 800,000 new cases diagnosed annually. Its major etiological factors include tobacco use, excessive alcohol consumption, and human papillomavirus (HPV) infection[1]. HNSCC exhibits profound molecular and clinical heterogeneity stratified by primary anatomical site and HPV status: HPV-associated oropharyngeal carcinomas typically confer a more favorable treatment response and prognosis, whereas HPV-negative tumors (e.g., oral squamous cell carcinoma, the focus of this study) are characterized by higher chromosomal instability and poorer clinical outcomes[2,3]. Despite the widespread adoption of multimodal therapy (surgery combined with chemoradiotherapy) and the modest survival benefit of immune checkpoint inhibitors in recurrent/metastatic disease, the 5-year overall survival rate of HNSCC patients remains stagnant at approximately 50%, with local recurrence and distant metastasis accounting for most treatment failures[4]. Conventional TNM staging, while informative for anatomical tumor extent, fails to capture intrinsic molecular features, resulting in divergent prognoses and suboptimal treatment intensity among patients with equivalent stages[5]. This clinical gap highlights an urgent need for robust molecular prognostic biomarkers and precise risk stratification, with metabolic dysregulation emerging as a promising investigative avenue.

Carnitine is a quaternary ammonium compound derived from lysine and methionine, whose core physiological role is to shuttle long-chain fatty acids across the mitochondrial inner membrane to initiate β-oxidation and energy production. This process is orchestrated by a set of carnitine-related genes (CRGs), including carnitine palmitoyltransferases (CPT1/CPT2) and carnitine-acylcarnitine translocase (SLC25A20) [6]. As a well-established hallmark of cancer, metabolic reprogramming allows tumor cells to utilize fatty acid oxidation in addition to aerobic glycolysis, particularly in nutrient-deprived or glycolysis-restricted microenvironments. Upregulation of the carnitine shuttle system sustains mitochondrial oxidative phosphorylation and NADPH generation, thereby alleviating oxidative stress and

promoting tumor cell survival[7]. Dysregulated CRG expression has been associated with tumor proliferation, invasion, and patient prognosis in multiple malignancies, including hepatocellular carcinoma, breast cancer, and colorectal cancer. Metabolomic studies have further documented prominent lipid metabolism dysfunction and altered serum free carnitine levels in head and neck tumors, implicating carnitine pathway dysregulation in HNSCC metabolic remodeling[8–11]. However, most existing investigations are limited to isolated functional validation of individual genes such as CPT1A. Systematic integrative analysis of CRGs as a functional gene set, along with assessment of their prognostic value in HNSCC, remains scarce, and their roles in tumor progression and microenvironment remodeling remain incompletely defined. Recent single-cell and functional studies indicate that metabolic reprogramming in HNSCC/OSCC extends beyond malignant cells to cancer-associated fibroblasts (CAFs), in which altered lipid metabolism can support tumor progression through metabolic and paracrine crosstalk[12,13]. This provides a rationale for tracing CRG dysregulation to fibroblast-state remodeling and for dissecting the links between metabolic features and tumor microenvironment remodeling at single-cell resolution.

Against this background, we analyzed transcriptomic data from the TCGA-HNSCC cohort. Differentially expressed genes between tumor and adjacent normal tissues were intersected with MSigDB carnitine metabolism gene sets to identify candidate CRGs specifically dysregulated in HNSCC. A six-gene prognostic risk score model was constructed using univariate and multivariate Cox regression with stepwise AIC selection, and its predictive performance was validated in the independent external cohort GSE41613. We further characterized the model's biological relevance and translational potential via multi-dimensional analyses, covering immune infiltration patterns, somatic mutation profiles, drug sensitivity prediction. Notably, integration of the single-cell RNA-seq dataset GSE181919 enabled cell-type assignment of prognostic genes and analysis of intercellular communication networks and metabolic pathway activity at single-cell resolution, revealing dynamic transcriptional and metabolic alterations of fibroblasts along differentiation trajectories in the tumor microenvironment—insights unattainable via bulk transcriptomics. Collectively, this study provides systematic evidence for the prognostic value of CRGs and their roles in HNSCC microenvironment remodeling at both bulk and single-cell levels, laying a foundation for future mechanistic investigations and targeted therapeutic development.

2. Results

2.1 Identification of DEGs and Carnitine-Related Candidate Genes and Their Functional Characterization in HNSCC

A total of 4114 DEGs ($|\log_2 FC| \geqslant 1$ and adjusted $P < 0.05$) were identified between tumor and normal tissues, including 2020 upregulated and 2094 downregulated genes (**Fig. 1a, Table S2**). Among the DEGs, HTN1, SMR3B, STATH and CST4 were markedly downregulated, whereas MAGEA12, MAGEA10, CT45A10 and MAGEC2 were markedly upregulated. Representative DEGs showed consistent expression separation between the Tumor and Normal groups (**Fig. 1b**). Intersection of the 4114 DEGs with 91 CRGs yielded 26 candidate genes (**Fig. 1c, Table S3**). THRSP, ACADM, ALDH9A1, ACACB and HADH were downregulated among the candidates, whereas MMP1, COL7A1 and SLC22A1 were upregulated, indicating that carnitine metabolism–related genes are broadly dysregulated in HNSCC.

Functional enrichment analysis of the 26 candidate genes ($P < 0.05$) identified 213 significant BP terms, 22 CC terms, 62 MF terms and 14 KEGG pathways (**Fig. 1d, Table S4**). For BPs, the candidate genes were enriched in carnitine metabolic process ($P = 2.60\times10^{-12}$) and amino-acid betaine transport ($P = 4.22\times10^{-12}$). For CCs, enrichment was observed for mitochondrial matrix ($P = 3.46\times10^{-3}$) and apical plasma membrane ($P = 1.52\times10^{-2}$). For MFs, the genes were enriched in phosphatidylcholine binding ($P = 2.41\times10^{-12}$) and acyl-CoA dehydrogenase activity ($P = 5.38\times10^{-7}$). KEGG analysis showed enrichment in fatty acid metabolism ($P = 7.29\times10^{-8}$), valine, leucine and isoleucine degradation ($P = 5.50\times10^{-6}$), and the PPAR signaling pathway ($P = 3.36\times10^{-5}$), indicating that the candidate genes are functionally concentrated in fatty acid β-oxidation and carnitine shuttle–related processes.

In the PPI network (STRING score ≥ 0.40) (**Fig. 1e**), ACADM, HADH, CPT1B, ACACB, SLC25A20, CRAT, ACADL, ALDH9A1, GATM, SHMT1, ACAD8, PRKAA2 and THRSP formed a core interaction module. Collectively, 4114 DEGs were

identified in HNSCC, and intersection with the CRG set yielded 26 candidate genes enriched in carnitine metabolism and fatty acid β-oxidation, with a tightly interconnected PPI module centered on ACADM, HADH, CPT1B, ACACB and SLC25A20, providing a basis for subsequent functional validation.

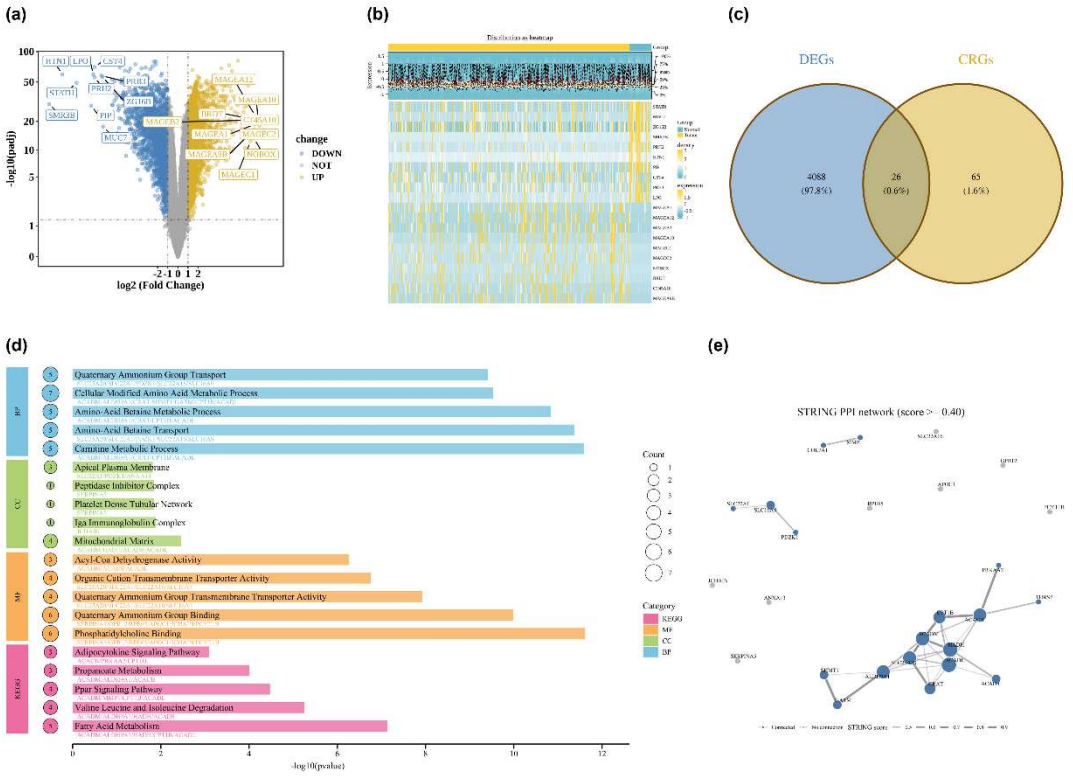


Fig. 1 Identification and characterization of candidate genes. (a) Volcano plot of DEGs between HNSCC tumor and normal tissues (adjusted $P < 0.05$, $|\log_2 FC| > 1$); yellow, upregulated; blue, downregulated. (b) Heatmap of representative DEGs across Tumor and Normal samples, with a density distribution panel above. (c) Venn diagram showing the overlap between DEGs and CRGs; the 26 intersecting genes were defined as candidate genes. (d) GO and KEGG enrichment dot-bar chart of the 26 candidate genes ($P < 0.05$); BP, biological process; CC, cellular component; MF, molecular function. (e) PPI network among candidate gene products (STRING, confidence score ≥ 0.40); node size reflects degree; edge width reflects interaction score

2.2 Identification of Six Carnitine-Related Prognostic Genes in HNSCC

Univariate Cox regression analysis of the 26 candidate genes identified eight candidate prognostic genes meeting the criteria of P < 0.2 and PH assumption test P > 0.05 (**Fig. 2a, Table S5**): ACADL (HR = 1.455, 80% CI: 1.219–1.736, P = 0.0067), SERPINA5 (HR = 1.293, 80% CI: 1.135–1.472, P = 0.0115), PRKAA2 (HR = 1.132, 80% CI: 1.042–1.229, P = 0.0552), ACACB (HR = 0.872, 80% CI: 0.773–0.983, P = 0.1423), SLC25A20 (HR = 0.820, 80% CI: 0.720–0.933, P = 0.0492), THRSP (HR = 0.716, 80% CI: 0.529–0.967, P = 0.155), CPT1B (HR = 0.539, 80% CI: 0.376–0.774, P = 0.0284) and GPR12 (HR = 0.211, 80% CI: 0.071–0.626, P = 0.0665), all satisfying the PH assumption (PH P > 0.05). Multivariate stepwise Cox regression retained six of these as the final prognostic genes (**Table 1**): ACADL, SERPINA5, PRKAA2, ACACB, SLC25A20 and GPR12, while THRSP and CPT1B were excluded from the final model. Based on the RiskScore calculated from the six prognostic genes, TCGA-HNSCC training-set patients were stratified at the maxstat-derived cutoff (−1.035) into 281 HRG and 168 LRG patients; the density of deaths increased markedly among patients with higher risk scores, and survival time shortened accordingly (**Fig. 2b**). Kaplan–Meier analysis showed that the HRG had significantly poorer OS than the LRG (**Fig. 2c**). Time-dependent receiver operating characteristic (ROC) analysis showed AUC values of 0.65, 0.66 and 0.65 for 1-, 3- and 5-year OS, respectively (**Fig. 2d**).

Table 1. Multivariate stepwise Cox regression analysis of the six carnitine metabolism-related prognostic genes in HNSCC

symbol	coef	exp(coef)	se(coef)	z	Pr(>|z|)
ACADL	0.4936	1.6382	0.1848	2.6708	0.0076*
SERPINA5	0.3320	1.3938	0.1030	3.2233	0.0012*
PRKAA2	0.1797	1.1969	0.1019	1.7633	0.0778
ACACB	-0.3793	0.6843	0.1491	-2.5430	0.0110*
SLC25A20	-0.2176	0.8045	0.1257	-1.7306	0.0835
GPR12	-1.8527	0.1568	0.9366	-1.9782	0.0479*

coef: regression coefficient; exp(coef): hazard ratio; se(coef): standard error of the coefficient; z: z-statistic; Pr(>|z|): p-value. *Genes with P < 0.05 were considered statistically significant

In the external validation cohort (GSE41613), RiskScore was calculated using the same

model coefficients, and patients were stratified at the validation-set cutoff (−4.470) into 10 HRG and 84 LRG patients, with higher risk scores again corresponding to a greater density of deaths and shorter survival time (**Fig. 2e**). The HRG showed significantly worse OS than the LRG (**Fig. 2f**), and time-dependent ROC analysis yielded AUC values of 0.70, 0.64 and 0.60 for 1-, 3- and 5-year OS, respectively (**Fig. 2g**). Collectively, the six-gene RiskScore model stratified HNSCC patients into groups with significantly different survival outcomes in both the training set and the independent external cohort (**Table S6**).

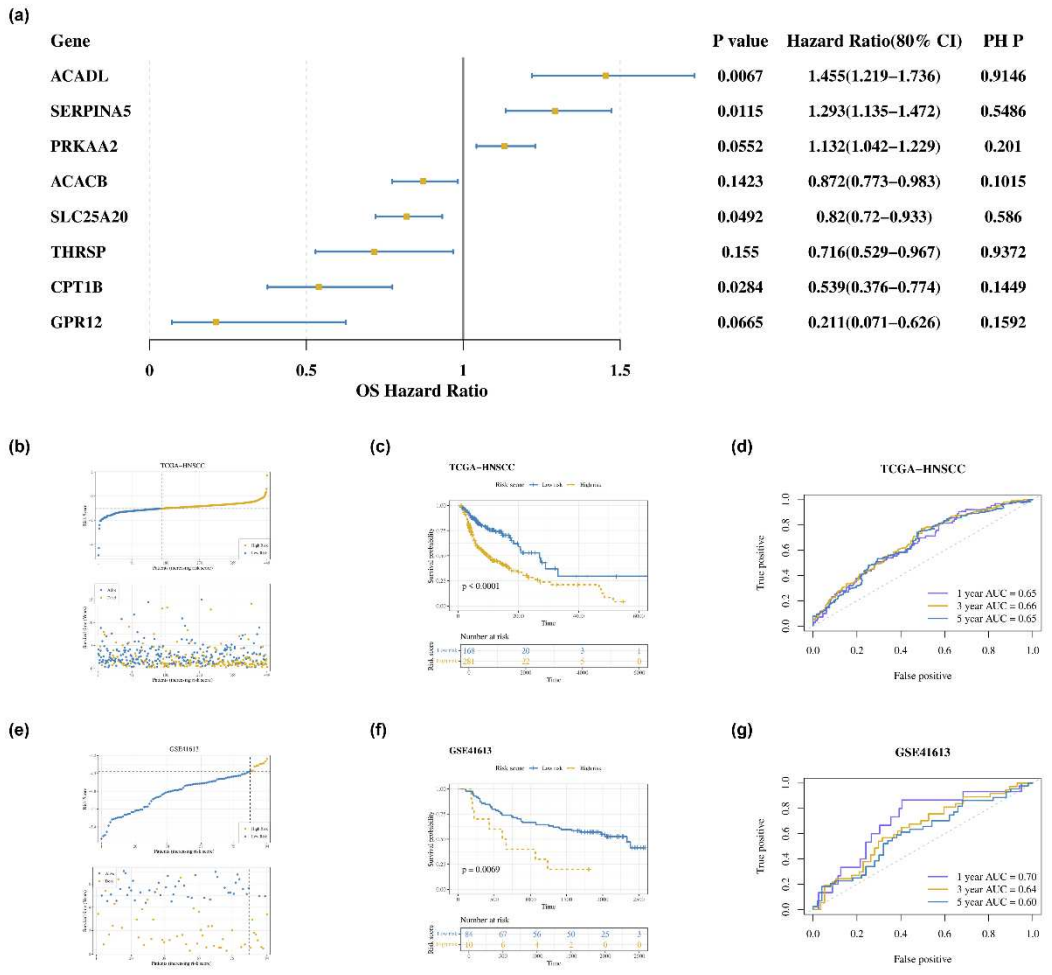


Fig. 2 Construction and validation of the prognostic risk model. (a) Forest plot of univariate Cox regression for candidate prognostic genes; PH, proportional hazards. (b) RiskScore distribution and survival status of TCGA-HNSCC training-set patients ranked by increasing risk score. (c) Kaplan–Meier survival curves of HRG and LRG in

143 the training set. (d) Time-dependent ROC curves for 1-, 3- and 5-year OS in the training
144 set. (e) RiskScore distribution and survival status of GSE41613 validation-cohort
145 patients ranked by increasing risk score. (f) Kaplan–Meier survival curves of HRG and
146 LRG in the GSE41613 cohort. (g) Time-dependent ROC curves for 1-, 3- and 5-year
147 OS in the GSE41613 cohort
148

149 **2.3 Construction of a RiskScore-Integrated Nomogram for Predicting Overall**
150 **Survival**

151 Univariate Cox regression combining RiskScore with clinicopathological variables
152 identified RiskScore (HR = 2.718, 95% CI 1.949–3.792, P = 3.89×10^{-9}), Age (HR =
153 1.019, 95% CI 1.006–1.032, P = 0.0044), T stage (T3, HR = 3.159, 95% CI 1.326–
154 7.525, P = 0.0094; T4, HR = 2.613, 95% CI 1.115–6.124, P = 0.0270), N stage (N2, HR
155 = 2.488, 95% CI 1.627–3.804, P < 0.001; N3, HR = 4.116, 95% CI 1.455–11.646, P =
156 0.0077), M stage (M1, HR = 132.996, 95% CI 8.319–2126.320, P < 0.001), pathologic
157 stage IV (P = 0.0447) and tumor grade G2 (P = 0.0304) as significant prognostic factors
158 (**Fig. 3a**). Multivariate stepwise Cox regression retained RiskScore, N stage, T stage
159 and Age in the final model; RiskScore remained an independent predictor of OS (P <
160 0.001), together with N stage (N2, P = 0.0033; N3, P = 0.0203), whereas T stage and
161 Age were no longer significant (**Fig. 3b**). Schoenfeld residual analysis showed no
162 significant time-dependent deviation for RiskScore, N stage, T stage or Age, confirming
163 that the proportional hazards assumption was satisfied for the multivariate model (**Fig.**
164 **3c, Table S7**).

165 A nomogram integrating RiskScore and N stage was constructed to predict 1-, 3- and
166 5-year OS probability (**Fig. 3d**). Calibration curves demonstrated close agreement
167 between predicted and observed OS probabilities for the 3- and 5-year models, while
168 the 1-year model showed a mild overestimation at lower predicted probabilities (**Fig.**
169 **3e**). Decision curve analysis indicated that the nomogram conferred clinically
170 meaningful net benefit over both the treat-all and treat-none strategies for 3- and 5-year

OS across a broad range of threshold probabilities (**Fig. 3f**). For 1-year OS, the nomogram avoided the net harm associated with the treat-all strategy at higher threshold probabilities, although its incremental benefit over the treat-none strategy was comparatively modest. Time-dependent ROC analysis of the nomogram yielded AUC values of 0.648, 0.736 and 0.713 for 1-, 3- and 5-year OS, respectively, exceeding the corresponding RiskScore-alone AUCs (**Fig. 3g**).

Collectively, the RiskScore-based nomogram combining RiskScore and N stage showed improved discrimination and clinical net benefit relative to RiskScore alone, supporting its potential utility for individualized OS prediction based on internal validation in the training set.

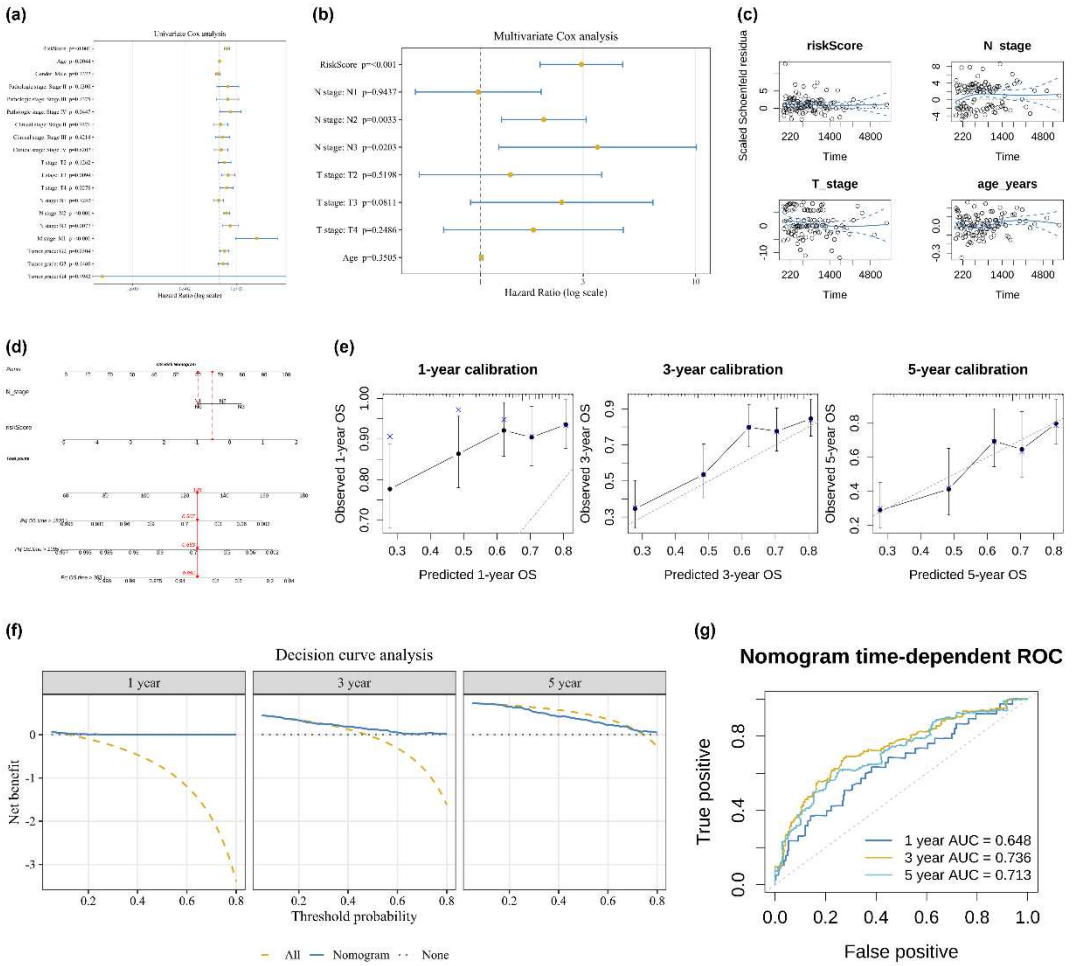


Fig. 3 Construction and evaluation of the prognostic nomogram. (a) Forest plot of univariate Cox regression for RiskScore and clinicopathological variables. (b) Forest

plot of multivariate Cox regression for variables retained after stepwise selection. (c) Scaled Schoenfeld residual plots testing the proportional hazards (PH) assumption for RiskScore, N stage, T stage and Age. (d) Nomogram predicting 1-, 3- and 5-year OS probability based on RiskScore and N stage. (e) Calibration curves for 1-, 3- and 5-year OS predicted by the nomogram. (f) Decision curve analysis (DCA) of the nomogram for 1-, 3- and 5-year OS. (g) Time-dependent ROC curves of the nomogram for 1-, 3- and 5-year OS

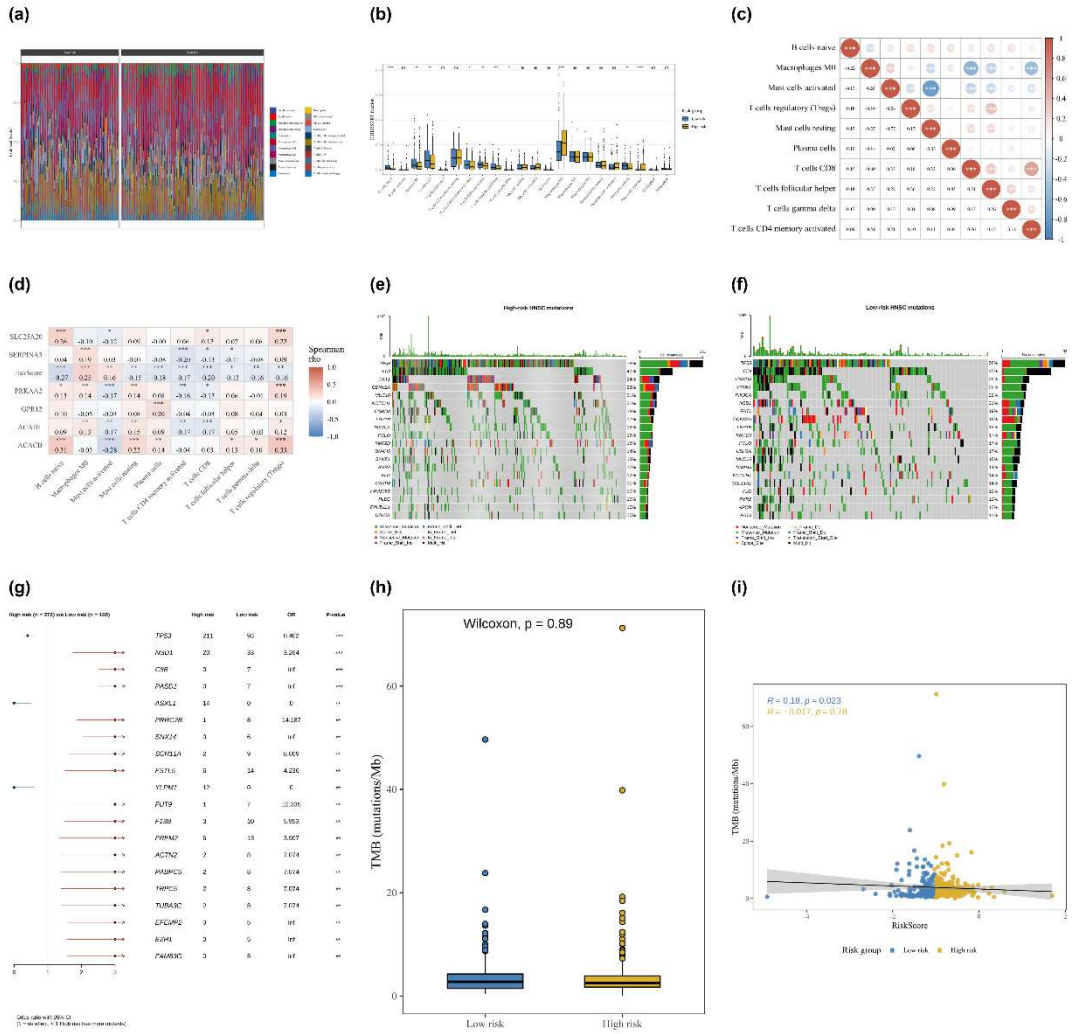
2.4 Differential Immune Cell Infiltration and Mutational Landscape between Risk Groups

CIBERSORT deconvolution showed that M0 macrophages, T cells CD4 memory resting and T cells CD8 were the predominant immune cell types in both risk groups (**Fig. 4a**). Wilcoxon rank-sum comparison identified significant differences between HRG and LRG for several immune cell types (**Fig. 4b**): M0 macrophages and mast cells activated were significantly increased in the HRG (both $P < 0.0001$), whereas T cells CD8, T cells follicular helper and T cells regulatory (Tregs) were significantly decreased in the HRG ($P < 0.01$, $P < 0.01$ and $P < 0.001$, respectively); the remaining cell types showed no significant difference.

Among immune cell pairs meeting the correlation threshold ($|cor| > 0.3$, $P < 0.05$), mast cells activated were negatively correlated with mast cells resting (cor = −0.775, $P = 4.17\times10^{-77}$), M0 macrophages were negatively correlated with T cells CD8 (cor = −0.599, $P = 2.39\times10^{-38}$) and T cells CD4 memory activated (cor = −0.513, $P = 7.03\times10^{-27}$), T cells CD8 were positively correlated with T cells CD4 memory activated (cor = 0.559, $P = 1.55\times10^{-32}$), T cells regulatory (Tregs) were positively correlated with T cells follicular helper (cor = 0.364, $P = 2.61\times10^{-13}$), and mast cells activated were negatively correlated with T cells CD8 (cor = −0.330, $P = 4.38\times10^{-11}$) (**Fig. 4c, Table S8**). For prognostic gene–immune cell correlations, only ACACB expression met the same threshold, showing positive correlations with B cells naive (cor = 0.306, P =

212 1.16×10^{-9}) and Tregs (cor = 0.327, P = 7.26×10^{-11}) (**Fig. 4d, Table S9**).

213 Oncoplots of the top 20 mutated genes showed TP53 as the most frequently mutated
214 gene in both the HRG (78%) and LRG (58%), followed by TTN (40% vs. 45%), FAT1
215 (24% vs. 19%) and CDKN2A (23% vs. 19%); NSD1 ranked among the top mutated
216 genes in the LRG (21%) but not in the HRG (**Fig. 4e–f**). Mutation frequency
217 comparison (HRG, n = 272; LRG, n = 160) showed that TP53 was significantly more
218 frequently mutated in the HRG (211/272, 78% vs. 93/160, 58%; P < 0.001), whereas
219 NSD1 and several genes including C8B, PASD1, PRRC2B, SNX14 and SCN11A were
220 significantly more frequently mutated in the LRG; ASXL1 and YLPM1 were mutated
221 exclusively in the HRG (all P < 0.05) (**Fig. 4g**).



222

223 **Fig. 4 Immune infiltration and mutational landscape of risk groups.** (a)

CIBERSORT-estimated immune cell composition in LRG and HRG. (b) Comparison of immune cell proportions between LRG and HRG. (c) Correlation matrix among immune cell types. (d) Correlation between prognostic gene expression and immune cell proportions. (e) Oncoplot of the top 20 mutated genes in the HRG. (f) Oncoplot of the top 20 mutated genes in the LRG. (g) Forest plot comparing mutation frequencies between HRG and LRG. (h) TMB comparison between LRG and HRG. (i) Correlation between RiskScore and TMB in LRG and HRG

TMB did not differ significantly between the LRG and HRG (Wilcoxon P = 0.89) (**Fig. 4h**). RiskScore was significantly positively correlated with TMB in the LRG (cor = 0.18, P = 0.023) but showed no significant correlation in the HRG (cor = −0.017, P = 0.78) (**Fig. 4i**). HRG and LRG showed distinct immune cell infiltration patterns and mutational profiles, with comparable overall tumor mutation burden between groups.

2.5 Risk-Associated and Prognostic Gene-Specific KEGG Pathway Enrichment

GSEA comparing HRG and LRG ($P < 0.05$, FDR < 0.25, |NES| > 1) showed that ribosome (NES = 2.93, $P = 2.19\times10^{-20}$) and oxidative phosphorylation (NES = 2.14, $P = 2.86\times10^{-9}$) were positively enriched in the HRG, whereas primary immunodeficiency (NES = −2.41, $P = 1.89\times10^{-10}$) and cell adhesion molecules (CAMs) (NES = −2.05, $P = 4.81\times10^{-8}$) were enriched in the LRG (**Fig. 5a, Fig. S1a**).

Single-gene GSEA of the prognostic genes identified largely distinct but partly overlapping pathway signatures relevant to tumor metabolism, adhesion, and cancer-associated signaling. ACACB expression was associated with cell adhesion molecules (NES = 1.44, $P = 3.81\times10^{-4}$) and the phosphatidylinositol signaling system (NES = 1.54, $P = 6.18\times10^{-4}$) (**Fig. 5b, Fig. S1b**). ACADL expression was associated with the neuroactive ligand-receptor interaction pathway (NES = 1.33, $P = 8.07\times10^{-5}$) and drug metabolism–cytochrome P450 (NES = 1.39, $P = 3.61\times10^{-3}$) (**Fig. 5c, Fig. S1c**). GPR12 expression was negatively associated with proteasome (NES = −3.57, $P = 2.36\times10^{-12}$) and positively associated with drug metabolism–cytochrome P450 (NES = 1.89, P =

252 3.25×10^{-9}) (**Fig. 5d, Fig. S1d**). PRKAA2 expression was associated with the insulin
253 signaling pathway (NES = 1.52, P = 1.20×10^{-5}) and the mTOR signaling pathway (NES
254 = 1.49, P = 5.63×10^{-3}) (**Fig. 5e, Fig. S1e**). SERPINA5 expression was associated with
255 ECM receptor interaction (NES = 2.10, P = 6.12×10^{-16}) and focal adhesion (NES = 1.80,
256 P = 1.72×10^{-16}) (**Fig. 5f, Fig. S1f**). SLC25A20 expression was associated with cell
257 adhesion molecules (NES = 1.93, P = 3.04×10^{-14}) and hematopoietic cell lineage (NES
258 = 1.92, P = 7.40×10^{-10}) (**Fig. 5g, Fig. S1g**).
259 Among the six prognostic genes, pathways in cancer was commonly and significantly
260 enriched in PRKAA2 (NES = 1.21, P = 5.54×10^{-3}), SERPINA5 (NES = 1.32, P =
261 5.02×10^{-5}) and SLC25A20 (NES = 1.22, P = 4.20×10^{-3}), suggesting a shared
262 association with oncogenic signaling among these prognostic genes. Collectively, HRG
263 and LRG showed distinct KEGG pathway enrichment at the group level, and each
264 prognostic gene showed a largely distinct pathway correlation profile, with pathways
265 in cancer as a shared feature among PRKAA2, SERPINA5 and SLC25A20(**Table S10**).

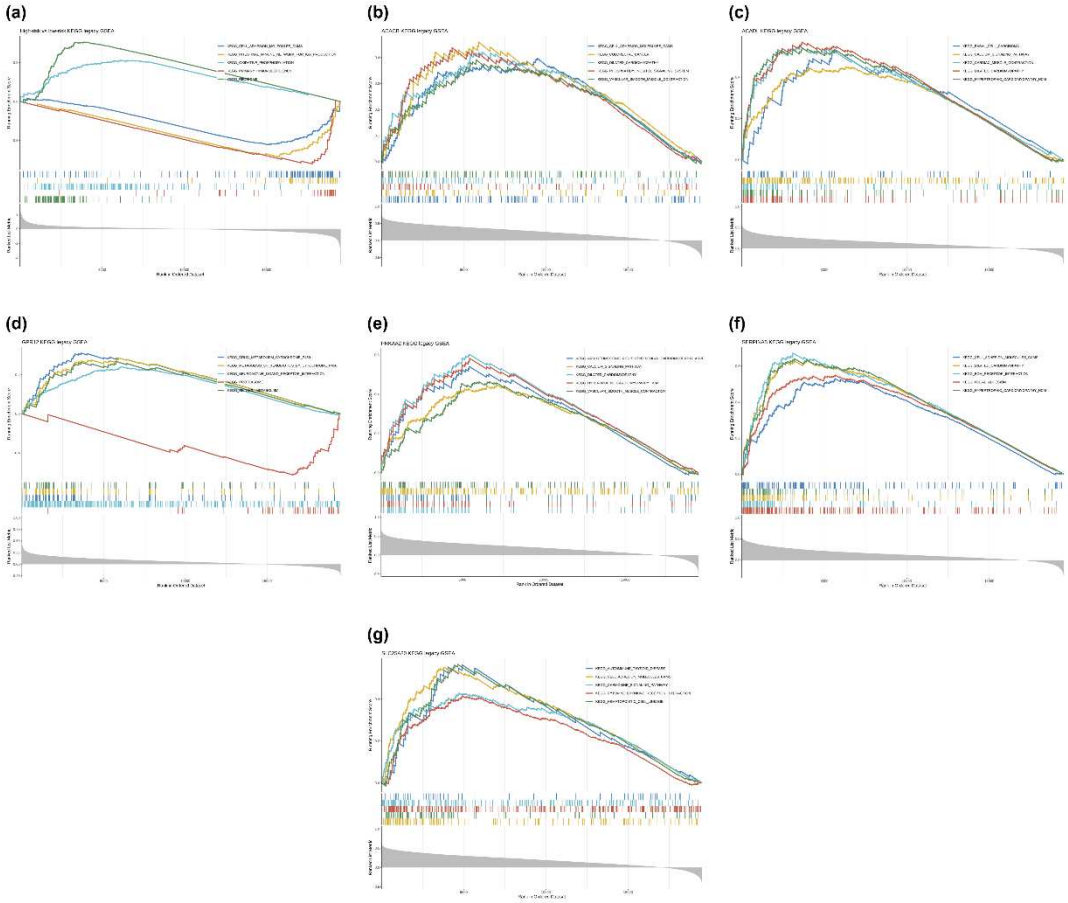


Fig. 5 GSEA of risk groups and prognostic genes. (a) Top KEGG legacy pathways enriched between HRG and LRG. (b) Top KEGG legacy pathways associated with ACACB expression. (c) Top KEGG legacy pathways associated with ACADL expression. (d) Top KEGG legacy pathways associated with GPR12 expression. (e) Top KEGG legacy pathways associated with PRKAA2 expression. (f) Top KEGG legacy pathways associated with SERPINA5 expression. (g) Top KEGG legacy pathways associated with SLC25A20 expression

2.6 Prognostic Genes Are Associated with Drug Sensitivity, Distinct Subcellular Localization and Functional Networks

Comparison of predicted IC_{50} between risk groups identified drugs with significantly higher predicted IC_{50} in the HRG, including AZD4547 ($P = 3.7\times10^{-25}$), PD173074 (P =

1.67×10^{-21}), Navitoclax ($P = 1.67\times10^{-20}$), EPZ004777 ($P = 1.7\times10^{-20}$) and Bortezomib ($P = 2.31\times10^{-19}$), indicating reduced predicted sensitivity to these agents in HRG patients (**Fig. 6a, Table S11**).

Correlation analysis between prognostic gene expression and predicted drug IC_{50} ($|cor| > 0.3$, $P < 0.05$) showed that ACACB expression was negatively correlated with IC_{50} of Axitinib ($cor = -0.735$, $P = 1.96\times10^{-77}$) and JQ1 ($cor = -0.732$, $P = 1.21\times10^{-76}$), SLC25A20 expression was negatively correlated with IC_{50} of Axitinib ($cor = -0.587$, $P = 5.41\times10^{-43}$), and PRKAA2 expression was positively correlated with IC_{50} of Obatoclax Mesylate ($cor = 0.577$, $P = 3.93\times10^{-41}$) (**Fig. 6b, Table S12**).

Subcellular localization analysis showed that ACADL, PRKAA2, ACACB and SLC25A20 were predominantly localized to the mitochondrion, whereas GPR12 was predominantly localized to the plasma membrane; all six genes showed detectable proportions in the nucleus and cytosol (**Fig. 6c**). The GeneMANIA functional network of the six prognostic genes and their interacting partners (including JPT2, ACACA, HADHA, STK11, PROC, THBD, FGG, FGB and CPT1B) was dominated by physical interactions, with annotated functions including regulation of coagulation, regulation of hemostasis, and fatty acid transmembrane transport (**Fig. 6d**). Collectively, the six prognostic genes showed distinct associations with predicted drug sensitivity, largely mitochondrial or membrane-associated subcellular localization, and a functional interaction network centered on coagulation- and lipid metabolism-related partners.

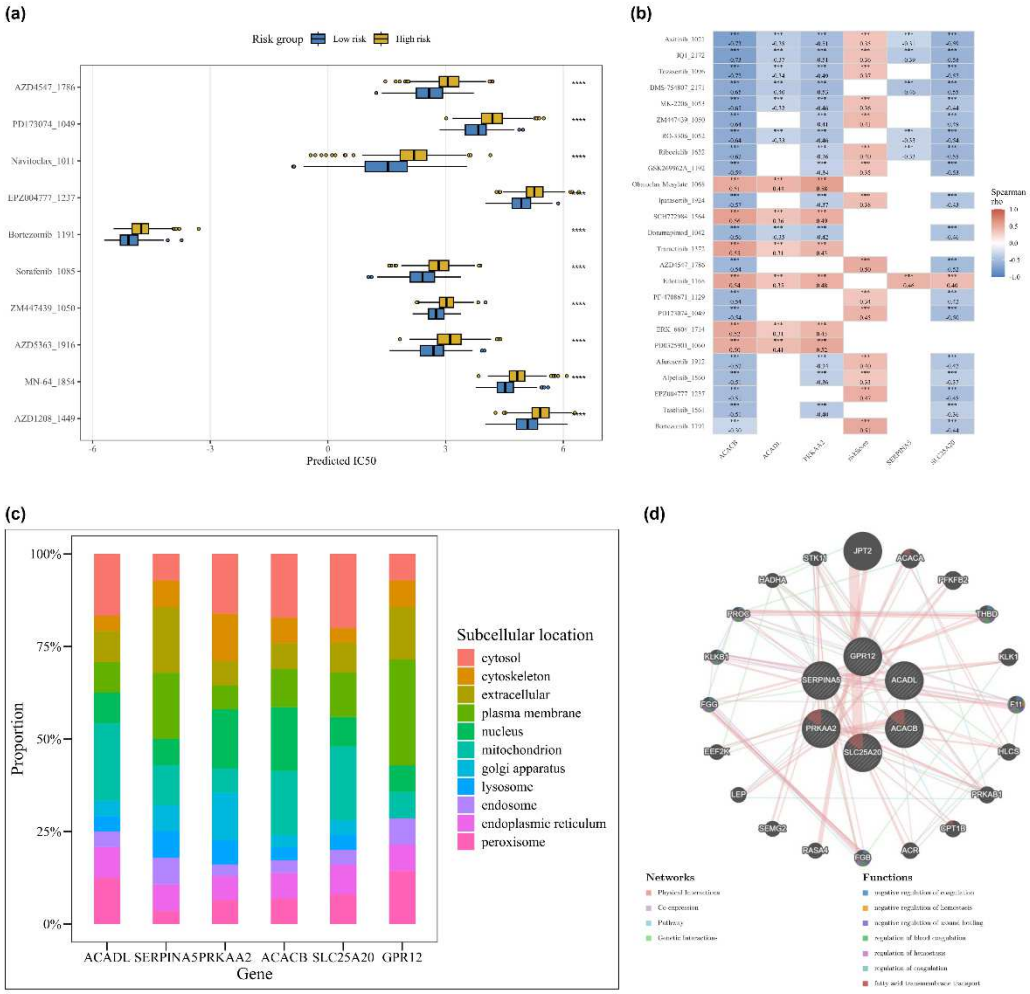


Fig. 6 Drug sensitivity, subcellular localization, functional network. (a) Predicted IC_{50} of the top 10 differential drugs between LRG and HRG. (b) Spearman correlation heatmap between prognostic gene expression and predicted drug IC_{50}. (c) Subcellular localization proportions of the six prognostic genes. (d) GeneMANIA functional interaction network of the prognostic genes

2.7 Single-Cell Landscape Identifies Fibroblasts as the Key Cell Population Linked to Prognostic Gene Expression

Quality-controlled cells (**Table S13**) were resolved into 26 preliminary clusters based on PCA of the top 2000 HVGs (**Fig. S2a**), with the elbow plot and JackStraw permutation test ($P < 0.05$) confirming the informativeness of the leading principal

311 components (**Fig. S2b–c**) and clustree analysis supporting a resolution of 0.2 (**Fig. S2d–**
312 **f**). These clusters were consolidated into nine major cell populations, which showed
313 clear separation between tumor and normal samples on UMAP (**Fig. 7a–b**).
314 Based on canonical markers (**Fig. 7c**), the nine populations were annotated as T cells
315 (XCL2, GZMH), Fibroblasts (PCOLCE2, TCF21), Macrophages (C1QC, FCN1), B
316 Plasma cells (BANK1, FCRL1), Epithelial cells (IVL, KRT19), Endothelial cells
317 (ACKR1, SELE), Dendritic cells (LILRA4, CLEC4C), Myocytes (MYOCD, CASQ2)
318 and Mast cells (TPSAB1, TPSB2). Cell composition differed markedly between groups
319 and was consistent across samples (**Fig. 7d, Fig. S2g**): Fibroblasts decreased in Tumor
320 ($P < 0.001$), while B Plasma cells ($P < 0.001$), Epithelial cells ($P < 0.01$), Macrophages
321 ($P < 0.01$), Dendritic cells ($P < 0.05$) and Mast cells ($P < 0.01$) increased; T cells showed
322 no difference (**Fig. 7e**).
323 Among the five prognostic genes detectable in the single-cell data, ACADL,
324 SERPINA5, PRKAA2, ACACB and SLC25A20 showed highest expression in
325 Fibroblasts, with PRKAA2 also enriched in Myocytes and ACACB in Endothelial cells
326 (**Fig. 7f**). Four genes (ACADL, PRKAA2, SERPINA5, SLC25A20) were significantly
327 differentially expressed in Fibroblasts between groups (all $P < 0.001$); ACACB was not
328 (**Fig. 7g**). GO BP enrichment of Fibroblast tumor-versus-normal DEGs (adjusted $P <$
329 0.05) was dominated by extracellular matrix organization, extracellular structure
330 organization and external encapsulating structure organization (all adjusted $P =$
331 5.57×10^{-21}), followed by collagen fibril organization (adjusted $P = 6.67\times10^{-19}$), collagen
332 metabolic process (adjusted $P = 9.15\times10^{-11}$) and cartilage development (adjusted $P =$
333 1.04×10^{-9}) (**Fig. 7h, Table S14**). These results identify Fibroblasts as the cell population
334 most consistently linked to tumor-associated compositional change and prognostic gene
335 dysregulation, with fibroblast-specific DEGs enriched in extracellular matrix-related
336 processes.

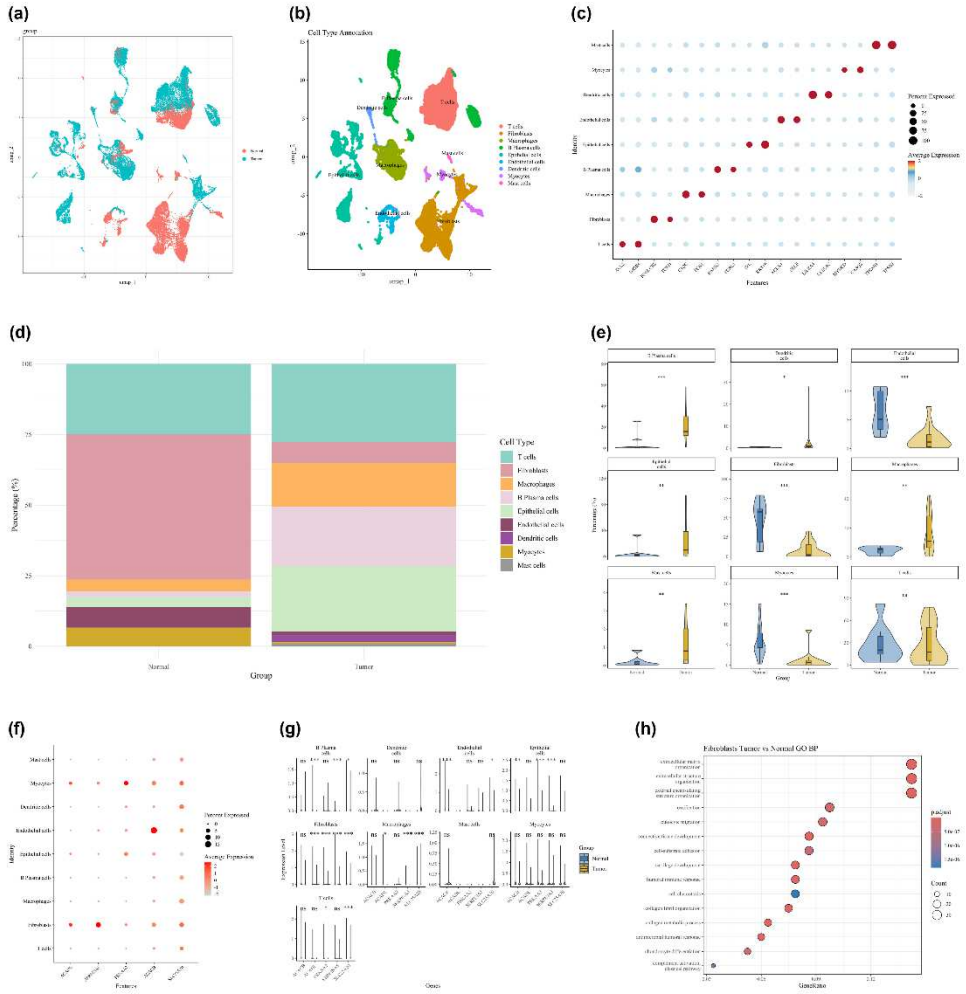


Fig. 7 Single-cell characterization of the tumor microenvironment. (a) UMAP plot colored by tumor/normal group. (b) UMAP plot colored by annotated cell type. (c) Dot plot of marker gene expression supporting cell type annotation. (d) Stacked bar plot of cell type proportions by group. (e) Violin plots comparing cell type proportions between tumor and normal groups. (f) Dot plot of prognostic gene expression across annotated cell types. (g) Violin plots of prognostic gene expression by cell type and group. (h) GO BP enrichment dot plot of tumor-vs-normal DEGs in Fibroblasts

2.8 Enhanced Fibroblast-Centered Intercellular Communication Network in Tumor Tissue

Cell–cell communication network analysis identified significant ligand–receptor pairs

among the nine cell populations in both groups. Tumor tissue showed higher inferred overall interaction number and strength than Normal (**Fig. S3a–b**). In both groups, Fibroblasts were the dominant hub, with the highest interaction count/strength as both sender and receiver, strong autocrine signaling, and prominent connections to Epithelial cells, Endothelial cells, Macrophages and Myocytes (**Fig. 8a–d**); this hub pattern was further strengthened in Tumor. With Fibroblasts as receiver, collagen (COL1A1/COL1A2/COL4A1/COL6A1-CD44), laminin (LAMA4/LAMB2/LAMC1-CD44), PDGFC/PDGFD-PDGFRA/PDGFRB, TGFB1-(TGFBR1+TGFBR2) and THBS1/THBS2-CD47 signals were prominent in Tumor but largely absent in Normal, mainly from Fibroblasts, Endothelial cells, Epithelial cells and Myocytes (**Fig. 8e**). As sender, Fibroblasts signaled to Endothelial cells, Epithelial cells, Macrophages and T cells via THY1-ADGRE5, THBS2-CD47, PDGFD-PDGFRB, NTN4-NTRK2 and CXCL12-CXCR4 in both groups; Tumor additionally showed expanded chemokine (CXCL1/2/6/8/13/16), complement (C3-C3AR1) and MHC-I/II (HLA-A/B/C, HLA-DRB1, HLA-DPB1)–CD4/CD8 signaling toward T cells, Dendritic cells and B Plasma cells (**Fig. 8f**). These results indicate that Fibroblasts showed the highest predicted interaction count and strength among the nine cell populations, with tumor-associated remodeling marked by expanded collagen/laminin–CD44, PDGF, TGFB, and THBS signaling.

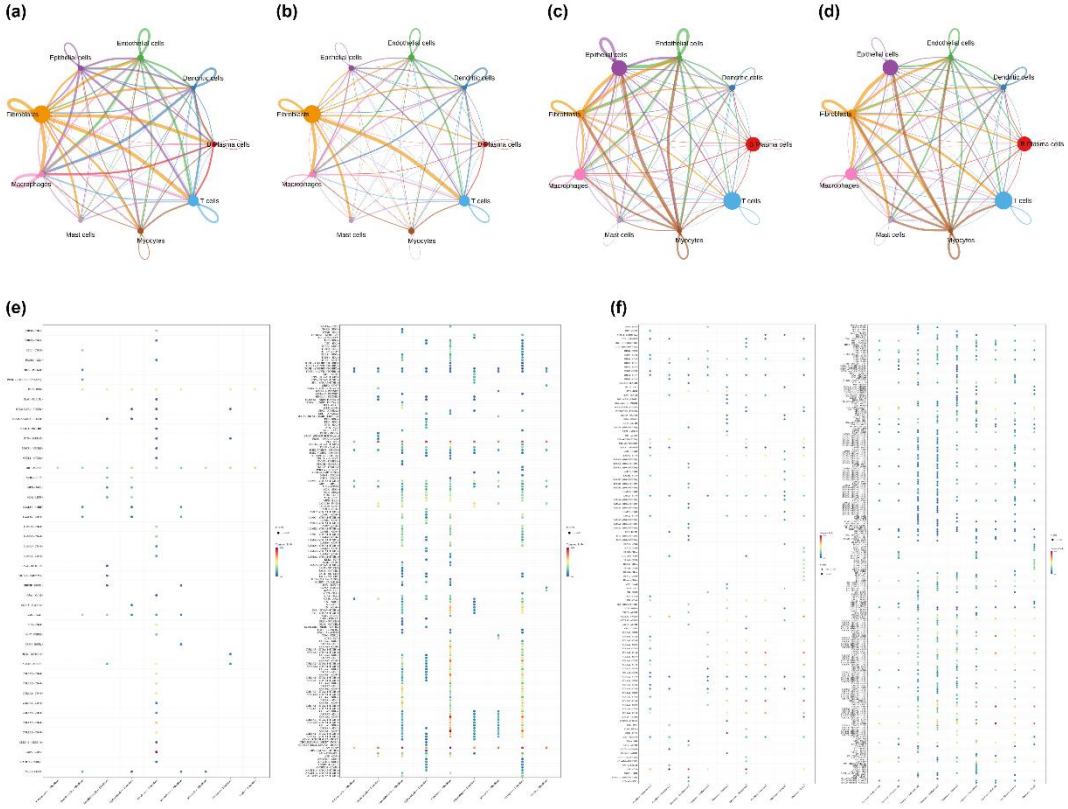


Fig. 8 Cell–cell communication network changes between Normal and Tumor tissue. (a) Circle plot of interaction number among nine cell types in Normal tissue. (b) Circle plot of interaction strength among nine cell types in Normal tissue. (c) Circle plot of interaction number among nine cell types in Tumor tissue. (d) Circle plot of interaction strength among nine cell types in Tumor tissue. (e) Bubble plot of significant ligand–receptor pairs with Fibroblasts as the receiving population, Normal (left) vs. Tumor (right). (f) Bubble plot of significant ligand–receptor pairs with Fibroblasts as the sending population, Normal (left) vs. Tumor (right)

2.9 Pseudotime Trajectory, Cell-Cycle Distribution and Metabolic Reprogramming of Key Cell Populations

Pseudotime analysis of the down-sampled Fibroblast subset resolved a branched trajectory with three states (State 1–3) rooted near the branch point. Fibroblasts_0 was Normal-enriched near the root at low pseudotime, while Fibroblasts_1 was Tumor-

383 enriched at the distal high-pseudotime branch (**Fig. 9a, Fig. S3c–e**), indicating a
384 pseudotime ordering from a normal-enriched state toward a tumor-enriched state under
385 the selected root. The five prognostic genes (ACACB, ACADL, PRKAA2, SERPINA5,
386 SLC25A20) were expressed mainly at low-to-intermediate pseudotime (Fibroblasts_0)
387 and largely lost at high pseudotime (Fibroblasts_1) (**Fig. 9b, Fig. 9c**), indicating
388 progressive down-regulation along this trajectory. Cell-cycle scoring showed G1-phase
389 predominance across all cell types and within Fibroblasts, with S/G2M cells forming
390 discrete peripheral subpopulations (**Fig. 9d–e**). Fibroblast S-phase proportion was
391 lower in Tumor than Normal, with a corresponding G1 increase (**Fig. 9f**).
392 Metabolic scoring showed cell-type-specific tumor-associated changes (**Fig. 9g**):
393 oxidative phosphorylation, valine/leucine/isoleucine biosynthesis and
394 phenylalanine/tyrosine/tryptophan biosynthesis were elevated in Tumor Epithelial cells
395 and Macrophages, while in Fibroblasts, metabolism of xenobiotics by cytochrome P450,
396 drug metabolism-cytochrome P450 and arginine biosynthesis were lower in Tumor.
397 Among the most significantly altered Fibroblast pathways, retinol metabolism, ether
398 lipid metabolism and fatty acid degradation were significantly decreased in Tumor (all
399 $P < 0.001$), while oxidative phosphorylation was significantly increased ($P < 0.001$)
400 (**Fig. S3f**). Collectively, these results suggest that the tumor-associated fibroblast state
401 is marked by loss of prognostic gene expression along pseudotime, reduced S-phase
402 proportion, and a shift from fatty-acid/xenobiotic metabolism toward oxidative
403 phosphorylation.

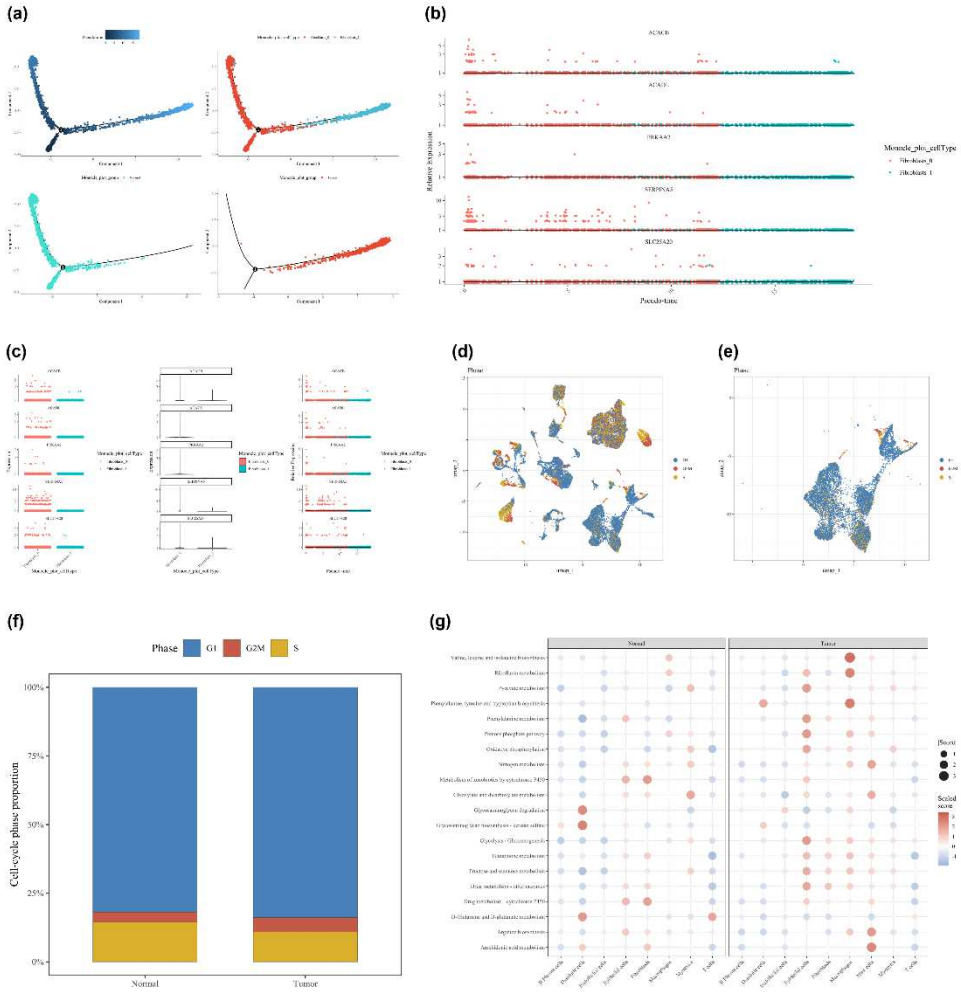


Fig. 9 Pseudotime trajectory, cell-cycle distribution, and metabolic activity of Fibroblasts. (a) Pseudotime trajectory of the down-sampled Fibroblast subset (DDRTree), colored by inferred pseudotime. (b) Heatmap of prognostic gene (ACACB, ACADL, PRKAA2, SERPINA5, SLC25A20) expression along the pseudotime trajectory. (c) Expression trends of the five prognostic genes across pseudotime, colored by inferred state. (d) UMAP plot of all annotated cell types colored by cell-cycle phase (G1, S, G2M). (e) UMAP plot of the Fibroblast subset colored by cell-cycle phase. (f) Cell-cycle phase proportions in Fibroblasts, Normal vs. Tumor. (g) Bubble plot of AUCell metabolic pathway activity scores across cell types, Normal vs. Tumor

3. Discussion

HNSCC displays marked molecular heterogeneity, and conventional TNM staging has inherent limits in stratifying survival disparities among patients with comparable anatomical stages, underscoring an unmet need for complementary molecular prognostic tools; yet the systematic prognostic relevance of carnitine metabolism—a core component of tumor lipid metabolic reprogramming—remains incompletely elucidated in HNSCC[5,14,15].

Using differential expression analysis, intersection with carnitine metabolism gene sets, and sequential univariate and multivariate Cox regression, we developed a six-gene prognostic risk model comprising *ACADL*, *SERPINA5*, *PRKAA2*, *ACACB*, *SLC25A20*, and *GPR12*, with 1-, 3-, and 5-year overall survival AUCs of 0.65, 0.66, and 0.65 in the training cohort, and externally validated it in the independent GSE41613 cohort. The high-risk subgroup was characterized by dysregulated immune infiltration, increased TP53 mutational frequency, and diminished predicted sensitivity to multiple agents including FGFR inhibitors and bortezomib; single-cell analysis further mapped expression alterations of five detectable prognostic genes to fibroblasts, which exhibited metabolic reprogramming (reduced fatty acid oxidation, elevated oxidative phosphorylation) along pseudotime trajectories and augmented hub status in the tumor intercellular communication network. Collectively, these findings provide systematic bulk and single-cell transcriptomic evidence delineating the prognostic value of carnitine metabolism-related genes in HNSCC and their roles in tumor microenvironment remodeling.

Biochemically, *ACACB*, *PRKAA2*, *SLC25A20*, and *ACADL* map to regulatory, transport, and oxidation components surrounding the canonical carnitine shuttle[6]. *ACACB*-encoded acetyl-CoA carboxylase β produces malonyl-CoA to allosterically inhibit CPT1 at the entry step[16], while *PRKAA2*-encoded AMPK α2 can relieve this inhibition by phosphorylating and suppressing ACC activity, thereby lowering malonyl-CoA, forming a canonical antagonistic regulatory pair[17]. *SLC25A20* mediates acylcarnitine translocation across the inner mitochondrial membrane[18], and *ACADL*

catalyzes the initial dehydrogenation step of long-chain fatty-acid β-oxidation[19], comprising the shuttle's transport-oxidation execution module.

The four-gene carnitine shuttle module showed discordant prognostic effects in univariate Cox analysis: *ACADL* (HR = 1.455) and *PRKAA2* (HR = 1.132) were risk factors, while *SLC25A20* (HR = 0.820) and *ACACB* (HR = 0.872) were protective, indicating non-uniform roles of distinct shuttle submodules in HNSCC progression, consistent with reported stage-dependent fatty acid oxidation utilization[20,21].

GSEA aligned with their molecular functions: *PRKAA2* in insulin/mTOR signaling[22]; *ACADL* in the KEGG basal cell carcinoma pathway (which should be interpreted as shared pathway-gene enrichment rather than histologic specificity[23,24]; and CYP450 drug metabolism; *SLC25A20* in cell adhesion and hematopoietic lineage pathways; *ACACB* in the cell adhesion pathway. Notably, *PRKAA2*, *SLC25A20*, and *SERPINA5* co-enriched in the pan-cancer "pathways in cancer" signature[25], suggesting their prognostic value extends beyond carnitine metabolism to shared oncogenic signaling.

At single-cell resolution, *ACADL*, *PRKAA2*, and *SLC25A20* are predominantly expressed in fibroblasts. *ACADL* and *SLC25A20* are markedly downregulated in tumor-resident fibroblasts, consistent with reduced fatty acid degradation and elevated oxidative phosphorylation, indicating carnitine shuttle dysregulation extends to stromal fibroblasts[26,27]. *PRKAA2* correlates positively with predicted IC_{50} of Bcl-2 inhibitor obatoclax mesylate, suggesting a potential association with AMPK-related drug resistance rather than chemosensitization[28]. *ACACB*, the sole immune-linked prognostic gene, correlates positively with Tregs and naïve B cells, implying shared upstream metabolic backgrounds without verified causality[29].

SERPINA5, a serine protease inhibitor family member that negatively regulates the coagulation-fibrinolysis axis, has been reported to modulate matrix degradation and tumor cell invasion by inhibiting extracellular protease activity[30]. In our study, *SERPINA5* was significantly enriched in ECM-receptor interaction and focal adhesion pathways, indicating its prognostic relevance is driven by tumor-ECM crosstalk rather than metabolic pathways[31]. Despite a predicted physical interaction with CPT1B in the GeneMANIA network, this association remains a suggestive network-level observation without biochemical validation[32].

GPR12, an orphan G protein-coupled receptor[33], exerts the strongest protective

477 effect (HR = 0.211) among the six prognostic genes included in our risk model. GSEA
478 showed its significant inverse association with the proteasome pathway and positive
479 association with CYP450-mediated drug metabolism, indicating statistical associations
480 with proteostasis- and xenobiotic-metabolism-related gene programs[34]. Nevertheless,
481 this remains a computational inference based on enrichment data. Its mechanistic role
482 in HNSCC is still undefined, as its endogenous ligands and downstream signaling
483 cascades have not been fully characterized.

484 Taken together, although both SERPINA5 and GPR12 are derived from the
485 intersection of carnitine-related gene sets and differentially expressed genes, their
486 prognostic associations map to two independent functional axes—ECM
487 remodeling/coagulation (SERPINA5) and protein homeostasis/xenobiotic metabolism
488 (GPR12)—rather than direct participation in carnitine shuttle transport or oxidation,
489 suggesting the six-gene risk model's prognostic performance arises from multiple
490 functionally distinct molecules rather than a single metabolic pathway.

491 Patients in the high-risk group exhibited significantly elevated proportions of M0
492 macrophages and activated mast cells, alongside reduced $CD8^+$ T cells, follicular helper
493 T (Tfh) cells, and regulatory T cells (Tregs), indicative of a tumor microenvironment
494 skewed toward innate immune infiltration and impaired adaptive antitumor immunity[35].
495 Increased M0 macrophages, an incompletely polarized subset, reflect dysregulated
496 local macrophage polarization rather than enrichment of immunosuppressive M2
497 macrophages—a nuance critical for immunological interpretation[36]. Concurrent
498 declines in $CD8^+$ T cells and Tfh cells align with the inferior survival of high-risk
499 patients, consistent with reports linking inadequate cytotoxic T-cell infiltration to
500 unfavorable HNSCC prognosis[37]. Recent spatial and single-cell profiling identified an
501 HNSCC CAF subset associated with restricted $CD8^+$ T-cell infiltration and antitumor
502 activity[38], providing a disease-specific context in which the reduced $CD8^+$ T-cell
503 abundance and enhanced fibroblast communication observed here may represent
504 coordinated features of an immune-excluded microenvironment rather than a
505 demonstrated causal axis.

506 Notably, decreased Treg abundance in the high-risk group runs counter to the
507 canonical paradigm that Treg accumulation fosters immune evasion and portends
508 poorer outcomes[39]. Nevertheless, the prognostic relevance of Tregs exhibits substantial
509 heterogeneity across malignancies; in select HNSCC cohorts, Treg infiltration

510 correlates with favorable survival, putatively by restraining aberrant inflammation and
511 preserving tissue homeostasis[40], supporting a context-dependent immunomodulatory
512 role of Tregs in the HNSCC microenvironment.

513 Somatic mutation profiling identified significant enrichment of *NSD1*, *C8B*,
514 *PASD1*, *PRRC2B*, *SNX14* and *SCN11A* mutations in the low-risk group. *NSD1*, a histone
515 methyltransferase whose loss-of-function mutations define a distinct HPV-negative
516 HNSCC subtype with unique methylation signatures and indolent clinical course[41],
517 showed enrichment concordant with favorable low-risk prognosis, suggesting
518 overlapping molecular traits and supporting combined stratification with metabolic risk
519 scores. Conversely, high-risk-exclusive *ASXL1* and *YLPM1* mutations, involved in
520 chromatin modification and transcriptional regulation[42], implicate epigenetic
521 dysregulation in the high-risk phenotype[43], though their low mutation frequency
522 warrants validation in larger independent cohorts.

523 Computational drug sensitivity profiling revealed diminished responsiveness to
524 multiple agents in the high-risk group, most notably FGFR inhibitors AZD4547 and
525 PD173074. As FGFR pathway activation is linked to targeted therapy response in
526 HNSCC subsets[44], this is consistent with lower predicted FGFR inhibitor sensitivity in
527 high-risk patients and provides a hypothesis for preclinical validation rather than direct
528 clinical stratification. Gene-level IC_{50} correlation analyses provided supporting
529 evidence at the gene level. *ACACB* and *SLC25A20* expression inversely correlated with
530 predicted IC_{50} of VEGFR inhibitor axitinib (cor = −0.735, −0.587), aligning with their
531 protective prognostic roles (HR = 0.872, 0.820) and indicating the low-risk state marked
532 by these genes confers favorable anti-angiogenic responses, supporting investigation of
533 metabolism-angiogenesis combination therapies[45]. *PRKAA2* expression positively
534 correlated with Bcl-2 inhibitor obatoclax mesylate IC_{50} (cor = 0.577), consistent with
535 AMPK-driven Bcl-2 inhibitor resistance via elevated autophagy[28], suggesting
536 autophagy inhibition may augment efficacy in AMPK-high subsets. Importantly, all
537 predictions represent computational inferences derived from GDSC2 cell line data and
538 expression profiles, not direct clinical drug response observations. Although these
539 predictions provide directional insights for drug screening and combinatorial regimens,
540 their translational utility warrants validation in organoid/cell-based assays and
541 prospective cohorts.

542 Single-cell analysis revealed a paradox among tumor-resident fibroblasts: despite

543 a lower captured proportion in the tumor scRNA-seq dataset, they were predicted to
544 form the core hub of the intercellular communication network. This pattern is consistent
545 with altered fibroblast state and communication activity rather than simple loss of
546 captured fibroblasts[46], concordant with the well-documented "low-abundance, high-
547 impact" paradigm of CAFs, which drive matrix remodeling and signaling via ECM
548 components and mediators across solid tumors[47]. Recent single-cell studies in HNSCC
549 identified functionally distinct CAF states, including SFRP2+ CAFs with strong
550 crosstalk with SPP1+ tumor-associated macrophages and pro-invasive CAF
551 populations[48,49], supporting the biological plausibility of a highly connected fibroblast
552 state despite limited captured representation; spatial and functional validation remains
553 necessary. Pseudotime trajectory analysis showed all five detectable prognostic genes
554 were highly expressed in normal-enriched Fibroblasts_0 (low-to-intermediate
555 pseudotime) but progressively downregulated in tumor-enriched Fibroblasts_1 (high
556 pseudotime). This downward trend is consistent with transcriptional differences
557 between normal-enriched and tumor-enriched fibroblast states, without establishing a
558 lineage transition[50]. Metabolically, CAFs exhibited diminished retinol, ether lipid and
559 fatty acid degradation alongside elevated oxidative phosphorylation, indicating
560 adaptive energy substrate reprogramming[51]. Concordant with downregulated fatty acid
561 oxidation genes (*ACADL*, *SLC25A20*), these findings are consistent with metabolic
562 remodeling in tumor-associated fibroblasts[52]. Experimental studies in OSCC have
563 shown enhanced mitochondrial oxidative phosphorylation in CAFs and tumor-driven
564 lipid biosynthesis in CAFs through the IL8/AKT/p-ACLY axis [53,13], indicating that the
565 elevated oxidative phosphorylation and reduced fatty-acid-degradation scores observed
566 here are biologically plausible features of CAF metabolic remodeling. Nevertheless,
567 because these scores reflect inferred pathway activity rather than direct metabolic flux,
568 they do not establish changes in metabolic flux or a true lineage transition between
569 normal-enriched and tumor-enriched fibroblast states; the proposed substrate
570 compensation mechanism therefore warrants validation via Seahorse assays or stable

571 isotope tracing.

572 Cell–cell communication analysis confirmed a markedly expanded fibroblast-
573 centered signaling network in tumors. Collagen/laminin–CD44, PDGF–PDGFR,
574 TGFB–TGFBR, and THBS–CD47 matrix-remodeling ligand–receptor pairs were
575 detected preferentially in tumors or showed higher inferred communication. These
576 well-validated drivers of fibroblast activation corroborate ECM-related differential
577 gene enrichment, offering convergent evidence from transcriptomic and cell–cell
578 communication dimensions[54]. As signal senders, tumor fibroblasts deliver expanded
579 signals to immune cells via chemokines, complement and MHC molecules, indicating
580 roles beyond structural support in immune recruitment and antigen presentation
581 modulation[55]. This lends transcriptomic support for immunoregulatory fibroblasts in
582 HNSCC, a subset characterized across multiple malignancies[56]. Experimental work in
583 OSCC has shown that CAF-intrinsic NNMT can promote angiogenesis and tumor
584 growth through epigenetic regulation of the ETS2/VEGFA axis[57], demonstrating that
585 CAF functions in HNSCC/OSCC can be established by mechanistic experiments rather
586 than inferred solely from ligand–receptor co-expression. This provides a functional
587 precedent for prioritizing the PDGF/TGFB/THBS and immune-related interactions
588 identified here for validation, while CellChat predictions themselves should not be
589 interpreted as confirmed signaling events.

590 Collectively, integrating TCGA-HNSCC transcriptomes with carnitine metabolism
591 gene sets, we developed and externally validated a six-gene prognostic risk model for
592 HNSCC. We systematically characterized the prognostic value of CRGs and their links
593 to tumor microenvironment remodeling, spanning immune infiltration, mutational
594 landscape, drug sensitivity, and single-cell fibroblast metabolic reprogramming. Our
595 findings offer a robust transcriptomic basis for downstream mechanistic studies.

596 Nevertheless, several limitations warrant acknowledgment. First, this study is
597 based entirely on retrospective, publicly available transcriptomic datasets of limited
598 size, including 449 TCGA-HNSCC samples for prognostic modeling, 94 HPV-negative

599 OSCC samples (GSE41613) for external validation, and 20 tumor plus 9 adjacent
600 normal samples (GSE181919) for single-cell analysis, and no experimental validation
601 was performed. Second, the external GSE41613 cohort exhibited pronounced sample
602 imbalance following maxstat stratification (10 vs. 84 patients), potentially
603 compromising the robustness of survival and ROC analyses; model generalizability
604 therefore requires replication in larger, well-balanced independent cohorts. Moreover,
605 the validation cohort was dichotomized using a validation-set-specific maxstat cutoff
606 rather than a prespecified training-derived threshold, which may introduce optimism
607 into survival separation and should be addressed in future validation. Third, immune
608 infiltration (CIBERSORT), drug sensitivity (oncoPredict), and CellChat-derived cell–
609 cell communication represent in silico transcriptomic predictions, denoting correlative
610 rather than causal associations. CellChat-identified ligand–receptor pairs mirror gene-
611 level coexpression, not validated protein-level signal transduction; all inferences
612 require experimental validation via immunohistochemistry, co-culture systems, or
613 functional perturbation assays. Fourth, low-resolution (0.02) re-clustering of
614 downsampled single-cell fibroblasts enabled pseudotime trajectory inference but
615 precluded identification of functional subsets (e.g., myofibroblast-like and
616 inflammatory cancer-associated fibroblasts); high-resolution fine-mapping of fibroblast
617 subpopulations remains to be addressed. Fifth, disease subtype heterogeneity across
618 training (full TCGA-HNSCC), validation (HPV-negative OSCC), and single-cell (no
619 explicit HPV stratification) cohorts may constrain cross-cohort generalizability of both
620 the model and single-cell observations.

621 Looking forward, functional validation of the fatty acid oxidation module (marked
622 by *ACADL*, *PRKAA2*, and *SLC25A20*) in HNSCC cell lines and organoid models should
623 be prioritized to delineate their causal contributions to tumor metabolic phenotypes and
624 tumor–fibroblast crosstalk. Concurrently, high-resolution CAF subtyping paired with
625 spatial transcriptomic mapping of key fibroblast subsets relative to tumor and immune
626 cells will enable finer mechanistic dissection of how fibroblast metabolic

627 reprogramming and signaling networks drive HNSCC progression.

628

629 # 4. Materials and methods

630 ## 4.1 Data sources

631 RNA sequencing expression data, somatic mutation data, and matched clinical and
632 survival information for HNSCC were obtained from The Cancer Genome Atlas
633 (TCGA, http://cancergenome.nih.gov/) database (TCGA-HNSC cohort; data accessed
634 on 30 May 2026). Samples included human primary tumor tissues and adjacent solid
635 normal tissues, yielding 515 samples (471 tumor and 44 normal) for differential
636 expression analysis between tumor and normal tissue. For prognostic modeling, a total
637 of 449 primary tumor samples were included after selecting cases with complete overall
638 survival (OS) status and follow-up time (OS.time) exceeding 90 days from the initial
639 471 tumor samples, thereby excluding those with insufficient follow-up[58].

640 An independent external validation cohort (GSE41613, platform GPL570) was
641 downloaded from the Gene Expression Omnibus (GEO,
642 https://www.ncbi.nlm.nih.gov/geo/) database (data accessed on 30 May 2026),
643 comprising human HPV-negative oral squamous cell carcinoma (OSCC) tissue. The
644 original cohort included 97 samples; after excluding samples with OS.time ≤ 90 days,
645 94 samples were retained for external validation of the risk model[59].

646 A single-cell RNA sequencing (scRNA-seq) dataset (GSE181919, platform GPL20301)
647 was obtained from the GEO database (data accessed on 30 May 2026), comprising 20
648 HNSCC tumor samples and 9 paired adjacent normal tissue samples, and was used for
649 single-cell-level analyses[60].

650 CRGs were retrieved from the Molecular Signatures Database (MSigDB,
651 https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) using the keyword “Carnitine”;
652 seven gene sets (GOBP_CARNITINE_METABOLIC_PROCESS,
653 GOBP_CARNITINE_TRANSPORT, REACTOME_CARNITINE_SHUTTLE,
654 GOBP_CARNITINE_TRANSMEMBRANE_TRANSPORT,

GOBP_CARNITINE_BIOSYNTHETIC_PROCESS, HP_ABNORMAL_CIRCULATING_CARNITINE_CONCENTRATION, and HP_DECREASED_PLASMA_CARNITINE) were merged and deduplicated to generate the CRG list used for subsequent intersection with differentially expressed genes (DEGs) (**Table S1**).

4.2 Differential expression analysis and functional characterization of candidate genes

Using the integer-count expression matrix of the TCGA-HNSCC training set (471 tumor and 44 normal samples, Section 4.1), differential expression analysis between tumor and normal tissue was performed with the R package "DESeq2" (v1.40.2) based on a negative binomial model[61]; genes with $|\log_2$ fold change ($\log_2$FC)$| \geq 1$ and adjusted $P < 0.05$ were defined as DEGs. Volcano plots were generated with the R packages "ggplot2" (v3.5.2) and "ggrepel" (v0.9.6, https://CRAN.R-project.org/package=ggrepel)[62], and heatmaps of DEGs were produced with the R package "ComplexHeatmap" (v2.16.0) [63].

DEGs were intersected with the CRG list described in Section 4.1, and the overlapping genes were defined as candidate genes; the overlap was visualized with a Venn diagram generated using the R package "ggvenn" (v0.1.10) [64].

Gene Ontology (GO; Biological Process (BP), Cellular Component (CC), and Molecular Function (MF)) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of candidate genes were performed with the R packages "clusterProfiler" (v4.10.1) and "org.Hs.eg.db" (v3.18.0, https://bioconductor.org/packages/org.Hs.eg.db/)[65], using a local MSigDB KEGG gene set file (Gene Matrix Transposed (GMT) format) as the pathway background and a significance threshold of $P < 0.05$; combined GO/KEGG bar plots were generated with "ggplot2" (v3.5.2), ranking pathways within each category by ascending P value. A protein–protein interaction (PPI) network of candidate genes was constructed via the application programming interface (API) of the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (https://string-db.org), restricted to *Homo sapiens* (NCBI Taxonomy ID: 9606) with an interaction confidence score threshold of 0.40; the network was visualized with the R packages "ggraph" (v2.2.1, https://CRAN.R-project.org/package=ggraph), "ggplot2" (v3.5.2), and "ggrepel"

(v0.9.6).

4.3 Prognostic gene screening, risk model construction and external validation

Among TCGA-HNSCC training-set tumor samples with complete OS information, univariate Cox regression analysis of candidate genes was performed with the R package "survival" (v3.8.3)[66]; genes with univariate Cox $P < 0.2$ and a proportional hazards (PH) assumption test $P > 0.05$ were retained as candidate prognostic genes. Candidate prognostic genes were entered into a multivariate Cox regression model, and bidirectional stepwise selection based on the Akaike Information Criterion (AIC) was performed with the stepAIC function of the R package "MASS" (v7.3.60.0.1) to identify the optimal model[67]; the genes retained in the final model were defined as prognostic genes. A risk score was calculated for each sample as the sum of the expression level of each prognostic gene weighted by its regression coefficient from the final Cox model:

$$risk = \sum_{(i=1)}^{n} Coef(i) \times expr(i)$$

where n denotes the number of prognostic genes retained in the final model, $Coef_i$ denotes the regression coefficient of the i-th prognostic gene derived from the final stepAIC Cox model fitted in the training set, and $Expr_i$ denotes the expression level of the i-th prognostic gene.

In the training set, the optimal RiskScore cutoff was determined using the maxstat method implemented in the R package "survminer" (v0.5.0)[68], and samples were stratified into a high-risk group (HRG) and a low-risk group (LRG) accordingly; Kaplan–Meier (KM) survival curves were plotted and compared using the log-rank test. Time-dependent receiver operating characteristic (ROC) curves and the area under the curve (AUC) at 1, 3, and 5 years were calculated with the R package "survivalROC" (v1.0.3.1) to evaluate the predictive performance of RiskScore for OS[69]. In the external validation cohort (GSE41613), RiskScore was calculated for each sample using the coefficients from the final training-set model, and the same maxstat method was applied to determine a validation-set-specific cutoff for risk stratification; KM and time-dependent ROC analyses were repeated as above to evaluate external validation performance.

4.4 Independent prognostic analysis and nomogram construction

RiskScore, calculated from the prognostic genes defined in Section 4.3, was integrated

719 with available clinical variables (Age, Gender, Pathologic stage, Clinical stage, T stage,
720 N stage, M stage, and Tumor grade) of TCGA-HNSCC training-set patients. Univariate
721 Cox regression combined with the PH assumption test was performed with the R
722 package "survival" (v3.8.3) to screen candidate variables; variables meeting the
723 univariate screening criteria were entered into a multivariate Cox regression model, and
724 those with an overall multivariate $P < 0.05$ and PH test $P > 0.05$ were retained for
725 nomogram construction. Nomogram models predicting 1-, 3-, and 5-year OS
726 probability were fitted with the R package "rms" (v6.8.1) and visualized with the R
727 package "regplot" (v1.1)[70,71]. Calibration curves based on 200 bootstrap resamples were
728 generated to evaluate agreement between predicted and observed probabilities, and
729 decision curve analysis (DCA) was performed to assess the clinical net benefit of the
730 nomogram model relative to the treat-all and treat-none strategies.

731 **4.5 Immune infiltration and mutational landscape analysis**

732 The relative proportions of 22 immune cell types in TCGA-HNSCC training-set tumor
733 samples were estimated with the CIBERSORT algorithm implemented in the R package
734 "IOBR" (v0.99.0, https://github.com/IOBR/IOBR); only samples with a CIBERSORT
735 deconvolution $P < 0.05$ were retained for group comparison[72]. Differences in immune
736 cell proportions between HRG and LRG were assessed with the Wilcoxon rank-sum
737 test. Correlations among immune cell proportions, and between RiskScore/prognostic
738 gene expression and immune cell proportions, were evaluated with Spearman
739 correlation analysis, using a threshold of |correlation coefficient (cor)| > 0.3 and $P <$
740 0.05.

741 Somatic mutation data for TCGA-HNSC were obtained with the R package
742 "TCGAmutations" (v0.3.0, https://bioconductor.org/packages/TCGAmutations/) and
743 matched to training-set risk groups by patient barcode using the R package "maftools"
744 (v2.18.0)[73]; oncoplots displaying the top 20 most frequently mutated genes were
745 generated separately for the HRG and LRG. Differences in mutation frequency between
746 HRG and LRG were assessed with the mafCompare function of "maftools" (v2.18.0),
747 and the top 20 genes showing the most significant mutation frequency differences ($P <$
748 0.05) were visualized as a forest plot. Tumor mutation burden (TMB) was calculated
749 based on a total exonic length of 38 Mb; differences in TMB between HRG and LRG

750 were assessed with the Wilcoxon rank-sum test, and the association between RiskScore
751 and TMB was evaluated with Spearman correlation analysis.

752 **4.6 Risk group and prognostic gene set enrichment analysis (GSEA)**

753 Based on HRG/LRG grouping of the TCGA-HNSCC training set, a genome-wide
754 ranked gene list was generated with the R package "DESeq2" (v1.40.2), and GSEA
755 against the MSigDB KEGG legacy gene sets was performed with the R package
756 "clusterProfiler" (v4.10.1), using significance thresholds of $P < 0.05$, false discovery
757 rate (FDR) < 0.25, and |normalized enrichment score (NES)| > 1; among the
758 significantly enriched pathways, the top 5 pathways ranked by significance (ascending
759 P value) were displayed. For each prognostic gene, the Spearman correlation coefficient
760 between its expression level and the genome-wide expression profile was calculated to
761 generate a gene-specific ranked list, and single-gene GSEA against the same MSigDB
762 KEGG legacy gene sets was performed using the same significance thresholds
763 described above; for each gene, the top 5 pathways were displayed as representative
764 enrichment curves.

765 **4.7 Drug sensitivity, subcellular localization, and functional network analysis of**
766 **prognostic genes**

767 Half-maximal inhibitory concentration (IC_{50}) values of candidate drugs for TCGA-
768 HNSCC training-set samples were predicted with the R package "oncoPredict" (v1.2)[74],
769 using expression and drug-response matrices from the GDSC2 dataset of the Genomics
770 of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) database.
771 Differences in predicted IC_{50} between HRG and LRG were assessed with the Wilcoxon
772 rank-sum test, and the top 10 drugs ranked by ascending P value were displayed as box
773 plots; associations between RiskScore/the prognostic genes identified in Section 4.3
774 and drug IC_{50} were evaluated with Spearman correlation analysis, using a threshold of
775 $|cor| > 0.3$ and $P < 0.05$.

776 Subcellular localization of the proteins encoded by the prognostic genes identified in
777 Section 4.3 was obtained from the GeneCards database (https://www.genecards.org/),
778 with localization evidence traced to the COMPARTMENTS database
779 (https://compartments.jensenlab.org/). Functional association networks of the

780 prognostic genes, including physical interaction, co-expression, pathway co-
781 membership, genetic interaction, and predicted relationships, were constructed using
782 the GeneMANIA database (http://genemania.org).

783 **4.8 Single-cell data quality control, clustering, and cell population annotation**

784 Processed unique molecular identifier (UMI) count matrices and barcode metadata for
785 the GSE181919 single-cell dataset were imported with the R package "Seurat"
786 (v5.2.99.9006)[75], retaining tumor and adjacent normal samples. The matrix followed
787 the quality-control criteria applied by the original data depositors (cells with 200–8000
788 detected genes and mitochondrial gene content below 10% of total UMI counts; genes
789 detected in more than 0.1% of cells), and no additional cell-level filtering was
790 performed. Data were normalized using the LogNormalize method, and the top 2000
791 highly variable genes (HVGs) were identified with the FindVariableFeatures function
792 (vst method). Principal component analysis (PCA) was performed with RunPCA (npcs
793 = 50), and the number of informative principal components was assessed using an
794 elbow plot and the JackStraw procedure. Based on the top 50 principal components, a
795 shared nearest-neighbor graph was constructed with FindNeighbors, dimensionality
796 reduction was performed with RunUMAP for Uniform Manifold Approximation and
797 Projection (UMAP), and cell clustering was performed with FindClusters (resolution =
798 0.2)[60].

799 Cell types were annotated based on cluster-specific marker genes identified by "Seurat"
800 (v5.2.99.9006), primarily referencing lineage marker signatures curated for HNSCC in
801 the literature[60], with HNSCC-relevant tissue markers from the CellMarker Human
802 database (v2.0, http://bio-bigdata.hrbmu.edu.cn/CellMarker/) used as supporting
803 evidence; final cell-type assignments were determined by integrating marker expression
804 evidence. UMAP plots were used to display the spatial distribution of annotated cell
805 types and tumor/normal grouping, and bubble plots were used to display representative
806 marker expression across clusters or cell types.

807 The proportion of each cell type was calculated per sample, and differences between

808 tumor and normal groups were assessed with the Wilcoxon rank-sum test; expression
809 differences of prognostic genes within each cell type between tumor and normal groups
810 were also compared. Cell types showing significant differences in proportion together
811 with consistent differential expression evidence for prognostic genes were defined as
812 key cell types. GO BP enrichment analysis of tumor-vs-normal differential marker
813 genes in key cell types was performed with the R package "clusterProfiler" (v4.10.1),
814 using a significance threshold of adjusted $P < 0.05$.

815 **4.9 Pseudotime, cell–cell communication, cell cycle, and metabolic pathway**
816 **analysis of key cell populations**

817 Key cell subsets were extracted and down-sampled to a maximum of 3000 cells per
818 group (tumor and normal); re-clustering of the down-sampled cells was performed with
819 the R package "Seurat" (v5.2.99.9006) at a resolution of 0.02. Pseudotime trajectories
820 were constructed with the R package "monocle" (v2.30.1) using 1000 ordering genes
821 and Discriminative Dimensionality Reduction via Learning a Tree (DDRTree)[76], with
822 the state containing the highest proportion of normal cells defined as the root state;
823 expression changes of prognostic genes and key cell subpopulations along pseudotime
824 were examined.

825 Cell–cell communication networks for the normal and tumor groups were constructed
826 separately with the R package "CellChat" (v2.2.0) using the CellChatDB.human
827 ligand–receptor database[77]; the number and strength of interactions between cell types
828 were compared, with emphasis on differential ligand–receptor interactions involving
829 key cell types as signal senders or receivers.

830 Cell-cycle scores were assigned to all annotated single cells using the built-in S-phase
831 and G2M-phase marker gene sets in "Seurat" (v5.2.99.9006, CellCycleScoring
832 function); the proportions of cells in G1, S, and G2M phases within key cell types were
833 calculated per sample, and differences between tumor and normal groups were
834 compared with the Wilcoxon rank-sum test.

835 Single-cell metabolic pathway activity was scored with the R package "scMetabolism"

836 (v0.2.1) combined with the R package "AUCell" (v1.22.0) using KEGG metabolism-
837 related gene sets[26,78]; metabolic pathway activity was compared between tumor and
838 normal groups, with emphasis on key cell types, using a significance threshold of $P <$
839 0.05. Heatmaps and bubble plots were used to display representative pathway activity
840 across cell types and groups, and violin plots were used to display the most significantly
841 different pathways in key cell types.

842 **4.10 Statistical analysis**

843 All statistical analyses were performed in R (v4.3.3). Continuous variables were
844 compared between two groups using the Wilcoxon rank-sum test, and survival
845 differences between groups were assessed using the log-rank test. Correlations between
846 continuous variables were evaluated with Spearman correlation analysis. Univariate
847 and multivariate Cox proportional hazards regression were used to evaluate
848 associations between variables and OS, and the PH assumption was tested for each Cox
849 model. Unless otherwise specified, a two-sided $P < 0.05$ was considered statistically
850 significant; analysis-specific thresholds (e.g., for DEG screening, GSEA, and
851 correlation analyses) are described in the corresponding sections above.

852 **Reference**

853 1. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors,
854 and Prevention of Head and Neck Squamous Cell Carcinoma. *Med Sci (Basel)*.
855 2023;11(2):42. doi:10.3390/medsci11020042

856 2. Chuang YH, Lin CY, Lee JC, et al. Identification of the HNSC88 Molecular
857 Signature for Predicting Subtypes of Head and Neck Cancer. *Int J Mol Sci*.
858 2023;24(17):13068. doi:10.3390/ijms241713068

859 3. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head
860 and neck squamous cell carcinoma. *Nat Rev Dis Primers*. 2020;6(1):92.
861 doi:10.1038/s41572-020-00224-3

862 4. Poulose JV, Kainickal CT. Immune checkpoint inhibitors in head and neck
863 squamous cell carcinoma: A systematic review of phase-3 clinical trials. *World J*
864 *Clin Oncol*. 2022;13(5):388-411. doi:10.5306/wjco.v13.i5.388

865 5. He S, Chen H, Li C, et al. Identification and validation of M2 macrophage-related

866 gene signature as a novel prognostic model for head and neck squamous cell
867 carcinoma. *Sci Rep*. 2024;14(1):25338. doi:10.1038/s41598-024-76866-0

868 6. Kim YA, Lee Y, Kim MS. Carnitine Shuttle and Ferroptosis in Cancer. *Antioxidants*
869 *(Basel)*. 2025;14(8):972. doi:10.3390/antiox14080972

870 7. Ma Y, Temkin SM, Hawkridge AM, et al. Fatty acid oxidation: An emerging facet
871 of metabolic transformation in cancer. *Cancer Lett*. 2018;435:92-100.
872 doi:10.1016/j.canlet.2018.08.006

873 8. Yuan P, Mu J, Wang Z, et al. Down-regulation of SLC25A20 promotes
874 hepatocellular carcinoma growth and metastasis through suppression of fatty-acid
875 oxidation. *Cell Death Dis*. 2021;12(4):361. doi:10.1038/s41419-021-03648-1

876 9. Jariwala N, Mehta GA, Bhatt V, et al. CPT1A and fatty acid β-oxidation are
877 essential for tumor cell growth and survival in hormone receptor-positive breast
878 cancer. *NAR Cancer*. 2021;3(3):zcab035. doi:10.1093/narcan/zcab035

879 10. Liu J, Li Y, Xiao Q, et al. Identification of CPT2 as a prognostic biomarker by
880 integrating the metabolism-associated gene signature in colorectal cancer. *BMC*
881 *Cancer*. 2022;22(1):1038. doi:10.1186/s12885-022-10126-0

882 11. Li W, Li Y, Chen Z, et al. Analysis of metabolomics and transcriptomics data in
883 head and neck squamous cell carcinoma. *Oral Dis*. 2023;29(4):1464-1479.
884 doi:10.1111/odi.14125

885 12. Meng X, Zheng Y, Zhang L, Liu P, Liu Z, He Y. Single-Cell Analyses Reveal the
886 Metabolic Heterogeneity and Plasticity of the Tumor Microenvironment during
887 Head and Neck Squamous Cell Carcinoma Progression. *Cancer Res*.
888 2024;84(15):2468-2483. doi:10.1158/0008-5472.CAN-23-1344

889 13. Liu P, Wang Y, Li X, et al. Enhanced lipid biosynthesis in oral squamous cell
890 carcinoma cancer-associated fibroblasts contributes to tumor progression: Role of
891 IL8/AKT/p-ACLY axis. *Cancer Sci*. 2024;115(5):1433-1445.
892 doi:10.1111/cas.16111

893 14. Liu S, Liu W, Ding Z, et al. Identification and validation of a novel tumor driver
894 gene signature for diagnosis and prognosis of head and neck squamous cell
895 carcinoma. *Front Mol Biosci*. 2022;9:912620. doi:10.3389/fmolb.2022.912620

896 15. Sun L, Wang X, Chen L, et al. CPT1A mediates chemoresistance in human
897 hypopharyngeal squamous cell carcinoma via ATG16L1-dependent cellular
898 autophagy. *Cell Insight*. 2023;2(6):100127. doi:10.1016/j.cellin.2023.100127

899 16. Yu Y, Nie Q, Wang Z, Di Y, Chen X, Ren K. Targeting acetyl-CoA carboxylase 1

900 for cancer therapy. *Front Pharmacol*. 2023;14:1129010.
901 doi:10.3389/fphar.2023.1129010

902 17. Galic S, Loh K, Murray-Segal L, Steinberg GR, Andrews ZB, Kemp BE. AMPK
903 signaling to acetyl-CoA carboxylase is required for fasting- and cold-induced
904 appetite but not thermogenesis. *Elife*. 2018;7:e32656. doi:10.7554/eLife.32656

905 18. Tonazzi A, Giangregorio N, Console L, Palmieri F, Indiveri C. The Mitochondrial
906 Carnitine Acyl-carnitine Carrier (SLC25A20): Molecular Mechanisms of Transport,
907 Role in Redox Sensing and Interaction with Drugs. *Biomolecules*. 2021;11(4):521.
908 doi:10.3390/biom11040521

909 19. Zhao X, Qin W, Jiang Y, et al. ACADL plays a tumor-suppressor role by targeting
910 Hippo/YAP signaling in hepatocellular carcinoma. *NPJ Precis Oncol*. 2020;4:7.
911 doi:10.1038/s41698-020-0111-4

912 20. Nicolini A, Ferrari P. Involvement of tumor immune microenvironment metabolic
913 reprogramming in colorectal cancer progression, immune escape, and response to
914 immunotherapy. *Front Immunol*. 2024;15:1353787.
915 doi:10.3389/fimmu.2024.1353787

916 21. Aiderus A, Black MA, Dunbier AK. Fatty acid oxidation is associated with
917 proliferation and prognosis in breast and other cancers. *BMC Cancer*.
918 2018;18(1):805. doi:10.1186/s12885-018-4626-9

919 22. Trefts E, Shaw RJ. AMPK: restoring metabolic homeostasis over space and time.
920 *Mol Cell*. 2021;81(18):3677-3690. doi:10.1016/j.molcel.2021.08.015

921 23. Suchors C, Kim J. Canonical Hedgehog Pathway and Noncanonical GLI
922 Transcription Factor Activation in Cancer. *Cells*. 2022;11(16):2523.
923 doi:10.3390/cells11162523

924 24. Reimand J, Isserlin R, Voisin V, et al. Pathway enrichment analysis and
925 visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap.
926 *Nat Protoc*. 2019;14(2):482-517. doi:10.1038/s41596-018-0103-9

927 25. Shorning BY, Dass MS, Smalley MJ, Pearson HB. The PI3K-AKT-mTOR Pathway
928 and Prostate Cancer: At the Crossroads of AR, MAPK, and WNT Signaling. *Int J*
929 *Mol Sci*. 2020;21(12):4507. doi:10.3390/ijms21124507

930 26. Qiu X, Lin Y, Li Q, et al. Integrated single-cell and bulk RNA sequencing reveals
931 prognostic and immunotherapy-associated myofibroblastic cancer-associated
932 fibroblast subtypes in head and neck squamous cell carcinoma. *Transl Cancer Res*.
933 2025;14(6):3730-3745. doi:10.21037/tcr-2025-649

27. Wei R, Song J, Pan H, Liu X, Gao J. CPT1C-positive cancer-associated fibroblast facilitates immunosuppression through promoting IL-6-induced M2-like phenotype of macrophage. *Oncoimmunology*. 2024;13(1):2352179. doi:10.1080/2162402X.2024.2352179

28. Grenier A, Poulain L, Mondesir J, et al. AMPK-PERK axis represses oxidative metabolism and enhances apoptotic priming of mitochondria in acute myeloid leukemia. *Cell Rep*. 2022;38(1):110197. doi:10.1016/j.celrep.2021.110197

29. McDonald-Hyman C, Aguilar EG, Compeer EB, et al. Acetyl-CoA carboxylase 1 inhibition increases Treg metabolism and graft-versus-host disease treatment efficacy via mitochondrial fusion. *J Clin Invest*. 2025;135(23):e182480. doi:10.1172/JCI182480

30. Fan M, Xiong X, Han L, et al. SERPINA5 promotes tumour cell proliferation by modulating the PI3K/AKT/mTOR signalling pathway in gastric cancer. *J Cell Mol Med*. 2022;26(18):4837-4846. doi:10.1111/jcmm.17514

31. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun*. 2020;11(1):5120. doi:10.1038/s41467-020-18794-x

32. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605-D612. doi:10.1093/nar/gkaa1074

33. Li H, Zhang J, Yu Y, et al. Structural insight into the constitutive activity of human orphan receptor GPR12. *Sci Bull (Beijing)*. 2023;68(1):95-104. doi:10.1016/j.scib.2022.12.023

34. Zhang C, Li J, Tang Q, Li L, Cao D. Targeting proteostasis for cancer therapy: current advances, challenges, and future perspectives. *Mol Cancer*. 2025;24(1):265. doi:10.1186/s12943-025-02472-x

35. Zhang MN, Bao H, Ji X. Development and validation of a novel palmitoylation-related prognostic signature in head and neck squamous cell carcinoma. *Transl Cancer Res*. 2025;14(12):8347-8364. doi:10.21037/tcr-2025-1228

36. Gao J, Liang Y, Wang L. Shaping Polarization Of Tumor-Associated Macrophages In Cancer Immunotherapy. *Front Immunol*. 2022;13:888713. doi:10.3389/fimmu.2022.888713

37. Borsetto D, Tomasoni M, Payne K, et al. Prognostic Significance of CD4+ and

968 CD8+ Tumor-Infiltrating Lymphocytes in Head and Neck Squamous Cell
969 Carcinoma: A Meta-Analysis. *Cancers (Basel)*. 2021;13(4):781.
970 doi:10.3390/cancers13040781

971 38. Li C, Guo H, Zhai P, et al. Spatial and Single-Cell Transcriptomics Reveal a Cancer-
972 Associated Fibroblast Subset in HNSCC That Restricts Infiltration and Antitumor
973 Activity of CD8+ T Cells. *Cancer Res*. 2024;84(2):258-275. doi:10.1158/0008-
974 5472.CAN-23-1448

975 39. Pan Y, Zhou H, Sun Z, et al. Regulatory T cells in solid tumor immunotherapy:
976 effect, mechanism and clinical application. *Cell Death Dis*. 2025;16(1):277.
977 doi:10.1038/s41419-025-07544-w

978 40. Seminerio I, Descamps G, Dupont S, et al. Infiltration of FoxP3+ Regulatory T
979 Cells is a Strong and Independent Prognostic Factor in Head and Neck Squamous
980 Cell Carcinoma. *Cancers (Basel)*. 2019;11(2):227. doi:10.3390/cancers11020227

981 41. Bui N, Huang JK, Bojorquez-Gomez A, et al. Disruption of NSD1 in Head and
982 Neck Cancer Promotes Favorable Chemotherapeutic Responses Linked to
983 Hypomethylation. *Mol Cancer Ther*. 2018;17(7):1585-1594. doi:10.1158/1535-
984 7163.MCT-17-0937

985 42. Yu H, Hong J, Shin DY, Lee CH. The role of ASXL1, SRSF2, and EZH2 mutations
986 in chromatin dysregulation of myelodysplastic neoplasia and acute myeloid
987 leukemia. *Leukemia*. 2025;39(10):2329-2339. doi:10.1038/s41375-025-02657-9

988 43. Nebbioso A, Tambaro FP, Dell'Aversana C, Altucci L. Cancer epigenetics: Moving
989 forward. *PLoS Genet*. 2018;14(6):e1007362. doi:10.1371/journal.pgen.1007362

990 44. Wang Z, Anderson KS. Therapeutic Targeting of FGFR Signaling in Head and Neck
991 Cancer. *Cancer J*. 2022;28(5):354-362. doi:10.1097/PPO.0000000000000615

992 45. Iwamoto H, Abe M, Yang Y, et al. Cancer Lipid Metabolism Confers
993 Antiangiogenic Drug Resistance. *Cell Metab*. 2018;28(1):104-117.e5.
994 doi:10.1016/j.cmet.2018.05.005

995 46. Lavie D, Ben-Shmuel A, Erez N, Scherz-Shouval R. Cancer-associated fibroblasts
996 in the single-cell era. *Nat Cancer*. 2022;3(7):793-807. doi:10.1038/s43018-022-
997 00411-z

998 47. Sahai E, Astsaturov I, Cukierman E, et al. A framework for advancing our
999 understanding of cancer-associated fibroblasts. *Nat Rev Cancer*. 2020;20(3):174-
1000 186. doi:10.1038/s41568-019-0238-1

1001 48. Wang Q, Zhao Y, Tan G, Ai J. Single cell analysis revealed SFRP2 cancer associated

1002 fibroblasts drive tumorigenesis in head and neck squamous cell carcinoma. *NPJ*
1003 *Precis Oncol*. 2024;8(1):228. doi:10.1038/s41698-024-00716-5

1004 49. Cai Z, Chen L, Chen S, Fang R, Chen X, Lei W. Single-cell RNA sequencing
1005 reveals pro-invasive cancer-associated fibroblasts in hypopharyngeal squamous
1006 cell carcinoma. *Cell Commun Signal*. 2023;21(1):292. doi:10.1186/s12964-023-
1007 01312-z

1008 50. Jain S, Rick JW, Joshi RS, et al. Single-cell RNA sequencing and spatial
1009 transcriptomics reveal cancer-associated fibroblasts in glioblastoma with
1010 protumoral effects. *J Clin Invest*. 2023;133(5):e147087. doi:10.1172/JCI147087

1011 51. Zhang Q, Ding L, Li J, et al. Single-cell RNA sequencing of OSCC primary tumors
1012 and lymph nodes reveals distinct origin and phenotype of fibroblasts. *Cancer Lett*.
1013 2024;600:217180. doi:10.1016/j.canlet.2024.217180

1014 52. Liang J, Li L, Li L, et al. Lipid metabolism reprogramming in head and neck cancer.
1015 *Front Oncol*. 2023;13:1271505. doi:10.3389/fonc.2023.1271505

1016 53. Xiao L, Hu Q, Peng Y, et al. TRAP1 suppresses oral squamous cell carcinoma
1017 progression by reducing oxidative phosphorylation metabolism of Cancer-
1018 associated fibroblasts. *BMC Cancer*. 2021;21(1):1329. doi:10.1186/s12885-021-
1019 09049-z

1020 54. Biffi G, Tuveson DA. Diversity and Biology of Cancer-Associated Fibroblasts.
1021 *Physiol Rev*. 2021;101(1):147-176. doi:10.1152/physrev.00048.2019

1022 55. Kieffer Y, Hocine HR, Gentric G, et al. Single-Cell Analysis Reveals Fibroblast
1023 Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov*.
1024 2020;10(9):1330-1351. doi:10.1158/2159-8290.CD-19-1384

1025 56. Elyada E, Bolisetty M, Laise P, et al. Cross-Species Single-Cell Analysis of
1026 Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated
1027 Fibroblasts. *Cancer Discov*. 2019;9(8):1102-1123. doi:10.1158/2159-8290.CD-19-
1028 0094

1029 57. Wang X, Zhao H, Luo X, et al. NNMT switches the proangiogenic phenotype of
1030 cancer-associated fibroblasts via epigenetically regulating ETS2/VEGFA axis.
1031 *Oncogene*. 2024;43(35):2647-2660. doi:10.1038/s41388-024-03112-2

1032 58. Chi H, Xie X, Yan Y, et al. Natural killer cell-related prognosis signature
1033 characterizes immune landscape and predicts prognosis of HNSCC. *Front Immunol*.
1034 2022;13:1018685. doi:10.3389/fimmu.2022.1018685

1035 59. Lohavanichbutr P, Méndez E, Holsinger FC, et al. A 13-gene signature prognostic

of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res*. 2013;19(5):1197-1203. doi:10.1158/1078-0432.CCR-12-2647

60. Choi JH, Lee BS, Jang JY, et al. Single-cell transcriptome profiling of the stepwise progression of head and neck cancer. *Nat Commun*. 2023;14(1):1055. doi:10.1038/s41467-023-36691-x

61. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8

62. Slowikowski K. ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2' [computer program]. Version 0.9.6. 2024. Available from: https://CRAN.R-project.org/package=ggrepel

63. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847-2849. doi:10.1093/bioinformatics/btw313

64. Yan L. ggvenn: Draw Venn Diagram by 'ggplot2' [computer program]. Version 0.1.10. 2023. Available from: https://CRAN.R-project.org/package=ggvenn

65. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021;2(3):100141. doi:10.1016/j.xinn.2021.100141

66. Therneau TM. A Package for Survival Analysis in R [computer program]. Version 3.8-3. 2024. Available from: https://CRAN.R-project.org/package=survival

67. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002. doi:10.1007/978-0-387-21706-2

68. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2' [computer program]. Version 0.5.0. 2024. Available from: https://CRAN.R-project.org/package=survminer

69. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56(2):337-344. doi:10.1111/j.0006-341x.2000.00337.x

70. Harrell FE Jr. rms: Regression Modeling Strategies [computer program]. Version 6.8-1. 2024. Available from: https://CRAN.R-project.org/package=rms

71. Marshall R. regplot: Enhanced Regression Nomogram Plot [computer program]. Version 1.1. 2020. Available from: https://CRAN.R-project.org/package=regplot

72. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457. doi:10.1038/nmeth.3337

73. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756. doi:10.1101/gr.239244.118

74. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021;22(6):bbab260. doi:10.1093/bib/bbab260

75. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411-420. doi:10.1038/nbt.4096

76. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979-982. doi:10.1038/nmeth.4402

77. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088. doi:10.1038/s41467-021-21246-9

78. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083-1086. doi:10.1038/nmeth.4463

Data availability

The datasets analyzed in this study are publicly available. The TCGA-HNSC RNA-sequencing, somatic mutation, and corresponding clinical data were obtained from The Cancer Genome Atlas via the Genomic Data Commons portal (project TCGA-HNSC: https://portal.gdc.cancer.gov/projects/TCGA-HNSC). The external validation dataset GSE41613 (platform GPL570; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41613) and the single-cell RNA-sequencing dataset GSE181919 (platform GPL20301; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181919) were downloaded from the Gene Expression Omnibus. The carnitine-related gene sets

(GOBP_CARNITINE_METABOLIC_PROCESS, GOBP_CARNITINE_TRANSPORT, REACTOME_CARNITINE_SHUTTLE, GOBP_CARNITINE_TRANSMEMBRANE_TRANSPORT, GOBP_CARNITINE_BIOSYNTHETIC_PROCESS, HP_ABNORMAL_CIRCULATING_CARNITINE_CONCENTRATION, and HP_DECREASED_PLASMA_CARNITINE) were retrieved from the Molecular Signatures Database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp). Protein–protein interaction data were obtained from the STRING database (https://string-db.org), drug sensitivity data from the GDSC2 dataset of the Genomics of Drug Sensitivity in Cancer database (https://www.cancerrxgene.org/downloads/bulk_download), subcellular localization information from GeneCards (https://www.genecards.org/) with evidence traced to the COMPARTMENTS database (https://compartments.jensenlab.org/), and functional association networks from GeneMANIA (http://genemania.org). Cell-type marker evidence was obtained from the CellMarker Human database v2.0 (http://bio-bigdata.hrbmu.edu.cn/CellMarker/). All data generated or analyzed during this study are included in this article and its supplementary files. The R scripts used for data processing and statistical analysis are available from the corresponding author upon reasonable request.

Funding

This research did not receive funding.

Author contributions

LH and BW conceived and designed the research study. LH performed the formal data analysis, curated the research data, validated the analysis results, supervised the entire project, and administered the research program. BW conducted literature retrieval, completed bioinformatics analysis, and visualized the research figures. SD collected and curated the clinical data from public databases, and participated in statistical analysis. SJ provided resource support for the study, as well as verified the relevant

research data. RX revised the manuscript, polished the academic language, and provided professional research guidance. All authors have read and agreed to the published version of the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Supplementary Figure legends

Fig. S1 Dot plots of KEGG legacy GSEA results. (a) Top 5 pathways enriched between HRG and LRG. (b) Top 10 pathways associated with ACACB expression. (c) Top 10 pathways associated with ACADL expression. (d) Top 10 pathways associated with GPR12 expression. (e) Top 10 pathways associated with PRKAA2 expression. (f) Top 10 pathways associated with SERPINA5 expression. (g) Top 10 pathways associated with SLC25A20 expression

Fig. S2 Single-cell quality control and clustering diagnostics. (a) HVG selection plot. (b) PCA elbow plot. (c) JackStraw plot. (d) Clustree plot across clustering resolutions. (e) UMAP dimplot of preliminary clusters at resolution 0.2. (f) Marker gene dot plot across preliminary clusters. (g) Stacked bar plot of cell type proportions by individual sample

Fig. S3 Comparison of overall intercellular communication strength between Normal and Tumor tissue, and additional characterization of the fibroblast pseudotime trajectory. (a) Heatmap of pairwise interaction strength among cell types, Normal vs. Tumor. (b) Bar chart of total interaction number and strength, Normal vs.

Tumor. (c) UMAP plot of the Fibroblasts_0/Fibroblasts_1 subclusters. (d) UMAP plot of the Fibroblast subset colored by tissue group (Normal vs. Tumor). (e) Pseudotime trajectory of the Fibroblast subset colored by inferred state (State 1–3). (f) Violin plots of the top differentially active AUCell metabolism scores (retinol metabolism, ether lipid metabolism, fatty acid degradation, oxidative phosphorylation) in Fibroblasts, Normal vs. Tumor

Supplementary Table legends

Table S1. List of CRGs.
Table S2. Results of DEG analysis between HNSCC and normal tissues.
Table S3. Annotated results of DEG–CRG intersection candidate genes.
Table S4. GO and KEGG enrichment analysis results of candidate genes.
Table S5. Univariate Cox regression results of candidate prognostic genes.
Table S6. RiskScore and risk-group assignment of patients in the training and validation cohorts.
Table S7. Significant terms from univariate Cox regression of RiskScore and clinicopathological variables.
Table S8. Spearman correlation coefficients and P values among immune cell proportions meeting the threshold |cor| > 0.3 and P < 0.05, as estimated by CIBERSORT.
Table S9. Correlation results between prognostic gene expression and immune cell proportions.
Table S10. GSEA results comparing HRG and LRG, and single-gene GSEA results for prognostic genes, against KEGG legacy gene sets.
Table S11. Wilcoxon test results for predicted IC_{50} between HRG and LRG.
Table S12. Representative Spearman correlation results between prognostic gene expression and predicted drug IC_{50}.
Table S13. Per-sample, per-cell-type summary of single-cell quality-control metrics (cell number, median nFeature_RNA, median nCount_RNA, median mitochondrial percentage) before and after quality-control filtering, by tissue group (Tumor/Normal) and sample.
Table S14. GO Biological Process enrichment results for tumor-versus-normal differentially expressed genes in fibroblasts.

Abbreviations

Abbreviation	Full term

AIC	Akaike Information Criterion
AMPK	AMP-activated protein kinase
AUC	area under the curve
BP	Biological Process
CAF	cancer-associated fibroblast
CAM	cell adhesion molecule
CC	Cellular Component
CRG	carnitine-related gene
DCA	decision curve analysis
DDRTree	Discriminative Dimensionality Reduction via Learning a Tree
DEG	differentially expressed gene
FDR	false discovery rate
GDSC	Genomics of Drug Sensitivity in Cancer
GEO	Gene Expression Omnibus
GO	Gene Ontology
GSEA	gene set enrichment analysis
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus
HRG	high-risk group
HVG	highly variable gene
IC_{50}	half-maximal inhibitory concentration
KEGG	Kyoto Encyclopedia of Genes and Genomes
KM	Kaplan–Meier
LRG	low-risk group
MF	Molecular Function

MSigDB	Molecular Signatures Database
NES	normalized enrichment score
OS	overall survival
OSCC	oral squamous cell carcinoma
PCA	principal component analysis
PH	proportional hazards
PPI	protein–protein interaction
ROC	receiver operating characteristic
scRNA-seq	single-cell RNA sequencing
TCGA	The Cancer Genome Atlas
Tfh	follicular helper T cell
TMB	tumor mutation burden
Treg	regulatory T cell
UMAP	Uniform Manifold Approximation and Projection
UMI	unique molecular identifier

1191

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx
- SupplementaryMaterials.rar