scientific reports

OPEN

Identification of a hypoxia-related gene prognostic signature in colorectal cancer based on bulk and single-cell RNA-seq

Yihuan Qiao[1,5], Xunliang Jiang[2,3,5], Yaoting Li[4], Ke Wang[2,3], Rujie Chen[2,3], Jun Liu[2], Yongtao Du[2,3], Li Sun[1✉] & Jipeng Li[3✉]

Colorectal cancer (CRC) is the most common and fatal tumor in the gastrointestinal system. Its incidence and mortality rate have increased in recent years. Hypoxia, a persistent physiological tumor feature, plays a vital role in CRC tumorigenesis, metastasis, and tumor microenvironment (TME). Therefore, we constructed a hypoxia-related gene (HRG) nomogram to predict overall survival (OS) and explored the role of HRGs in the CRC TME. The Cancer Genome Atlas (TCGA) dataset was used as the training set, and two Gene Expression Omnibus datasets (GSE39582 and GSE103479) were used as the testing sets. HRGs were identified using the Gene Set Enrichment Analysis (GSEA) database. An HRG prognostic model was constructed in the training set using the least absolute shrinkage and selection operator regression algorithm and validated in the testing sets. Then, we analyzed tumor-infiltrating cells (TICs) using the cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm. Furthermore, single-cell next-generation RNA sequencing (RNA-seq) was used to investigate HRG expression in different TICs in the GSE139555 dataset. Finally, reverse transcription polymerase chain reactions (RT-PCR) were used to validate HRG mRNA expression in ten pairs of CRC normal and cancer tissue samples. A six HRG prognostic signature was constructed, with a superior OS prediction ability in CRC patients (area under the receiver operating characteristic curve (AUC) at one year: 0.693, AUC at three years: 0.712, and AUC at five years: 0.780). GSEA enrichment analysis identified six pathways enriched in the high-risk group. The TIC analysis indicated that the high-risk group had lower T-cell expression and higher neutrophil expression than the low-risk group. Furthermore, immune-related genes had an inseparable relationship with the HRG prognostic signature. Based on single-cell RNA-seq data, we found elevated hexokinase 1 (*HK1*) and glucose-6-phosphate isomerase (*GPI*) gene expression in natural killer (NK) and CD8[+] T cells. RT-PCR in ten CRC normal-tumor tissue pairs showed that expression of the signature's six HRGs varied differently in cancerous and paracancerous tissues. The constructed HRG signature successfully predicted the OS of whole-stage CRC patients. In addition, we showed that the signature's six HRGs were closely associated with the TME in CRC, where hypoxia inhibits the antitumor function of T cells.

Colorectal cancer (CRC) is the most commonly diagnosed gastrointestinal tumor, with the third highest incidence and second highest mortality[1]. CRC includes colon adenocarcinoma (COAD) and rectal adenocarcinoma. Due to delayed diagnosis, rapid progression, and early metastasis, CRC patients' overall survival (OS) is far from clinically satisfactory, even with recent dramatic advances in CRC treatment strategies[2,3]. Therefore, a superior prediction signature is urgently needed to ameliorate the current CRC diagnosis and treatment situation.

An important feature of solid tumors is hypoxia (low oxygen levels), which contributes to poor prognosis and therapeutic outcomes. Numerous studies have shown that hypoxia plays a crucial role in various critical cancer aspects, including angiogenesis[4], genome instability[5], metabolic reprogramming[6], epithelial-mesenchymal

[1]Department of Digestive Surgery, Honghui Hospital, Xi'an Jiaotong University, Xi'an 710054, Shaanxi, China. [2]State Key Laboratory of Cancer Biology, Department of Biochemistry and Molecular Biology, Air Force Medical University, Xi'an 710032, Shaanxi, China. [3]Department of Gastrointestinal Surgery, The First Affiliated Hospital of Air Force Medical University, Xi'an 710032, Shaanxi, China. [4]Xi'an Gaoxin No. 1 High School, Xi'an 710119, Shaanxi, China. [5]These authors contributed equally: Yihuan Qiao and Xunliang Jiang. ✉email: 364477087@qq.com; jipengli1974@aliyun.com

transition (EMT)[7], immune evasion[8], and therapy resistance[9]. Hypoxia induces increased expression of related genes and the creation of intratumoral oxygen gradients, such as reactive oxygen species (ROS), resulting in tumor plasticity and heterogeneity and enhanced tumor invasion and metastasis[10].

Hypoxia drives cancer progression mainly through hypoxia-inducible factors (HIFs). During normoxia, HIF-1α is hydroxylated and binds to the von Hippel-Lindau tumor suppressor (VHL) for ubiquitination[11]. HIF-1α rapidly accumulates in cells and drives gene transactivation during hypoxia. HIF-1α facilitates the glycolytic process by transactivating the expression of key glycolysis signaling pathway enzymes, including hexokinase 2 (HK2) and pyruvate kinase muscle isozyme M2 (PKM2)[12]. While HIF-1α is highly sensitive to oxygen concentration, HIF-1β is constitutively expressed regardless of the oxygen concentration[11]. The HIF-1α/HIF-1β complex can translocate to the nucleus and modulate genes with hypoxia-response consensus sequences[13]. In addition, HIF-1 inhibition by meloxicam had an anticancerous effect on hepatocellular carcinoma (HCC), contributing to caspase-dependent apoptosis of HCC cells during hypoxia[14]. In contrast to universal HIF-1α expression in almost all cells, HIF-2α is selectively expressed in tumor stem and endothelial cells. However, the underlying direct or indirect interaction mechanisms between hypoxia and CRC remain unclear.

Here, we evaluated the interaction between hypoxia-related genes (HRGs) and CRC, which could provide prognostic information for CRC patients. We developed a novel hypoxia gene signature independent of the current clinicopathologic characteristics and staging system through a series of systematic analyses to improve CRC prognostication. The study workflow is depicted in Fig. 1.

Materials and methods

Data acquisition. CRC patients' gene expression and corresponding clinicopathological features were obtained for the CRC cohort in The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) datasets GSE39582 and GSE103479. A single-cell next-generation RNA sequencing (RNA-seq) dataset for CRC patients was also obtained from the GEO database (GSE139555)[15]. The datasets generated and/or analysed during the current study are available in the TCGA and GEO repository, https://portal.gdc.cancer.gov/, https://www.ncbi.nlm.nih.gov/geo/.

Identification of differentially expressed and prognostic genes. Fifty-one HRGs were extracted from Gene Set Enrichment Analysis (GSEA) database[16]. The linear models for microarray data (*LIMMA*) R package was used to identify differentially expressed genes (DEGs) between paracancerous and cancerous samples based on a $P < 0.05$ and $|log_2$ fold change$| > 2$. OS-related DEGs were identified using a univariable Cox regression[17].
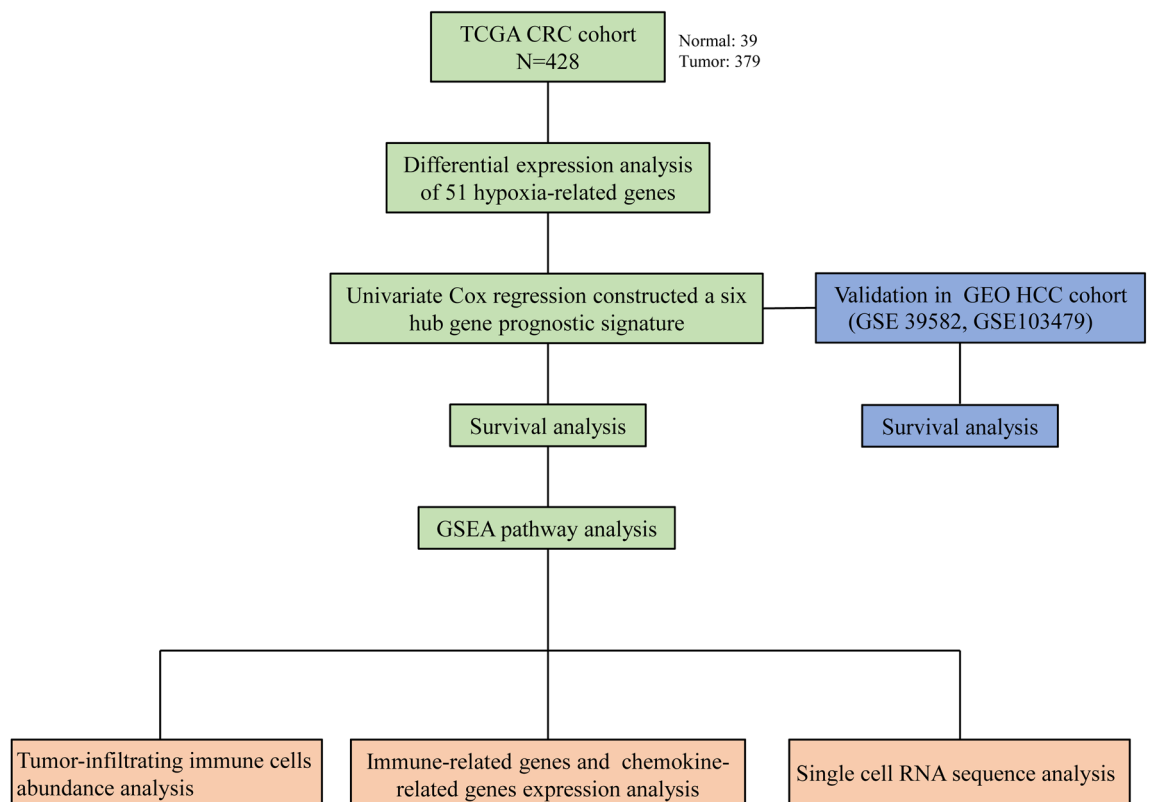
Figure 1. The study workflow.

HRG prognostic signature construction. The most predictive signature was constructed using multivariate Cox regression analyses with the *survival* R package. CRC patients were divided into two groups based on their median risk scores: high-risk and low-risk. Receiver operating characteristic (ROC) curves and Kaplan–Meier (K–M) curves were used to evaluate the risk model's predictive capability[18]. In addition, univariate and multivariate Cox regression analyses were used to explore the risk score model's prognostic efficiency independent of other clinicopathological features[17].

Functional enrichment analysis. GSEA was performed using the GSEA software (version 4.0.1) with 1000 permutation numbers. The cut-off criteria were gene size ≥ 15, |normalized enrichment score (NES)| > 1.5, and nominal $P < 0.05$.

Tumor-infiltrating cell analysis. The abundances of 22 tumor-infiltrating cell (TIC) types were estimated using the cell type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm[19] in the TCGA cohort, identifying significant results based on a $P < 0.05$.

Immune-related gene analysis. We analyzed the expression of 55 immune-related genes and nine chemokine-related genes extracted from the GSEA database to further explore the potential relationship between risk score and tumor immunity.

Gene expression quantification. Ten paired CRC tumor-normal tissues were obtained from patients in the Xijing Hospital (Xian, China). Informed consent was taken from all the patients as previously described[20]. Messenger RNA (mRNA) was extracted from each tissue using the TRIzol method (Invitrogen, Carlsbad, CA, USA). Quantitative real-time reverse transcription polymerase chain reactions (RT-PCR) were performed using SYBR-green PCR MasterMix (TaKaRa, Tokyo, Japan). Information related to the RT-PCR is listed in Table S1.

Single-cell RNA-seq and statistical analyses. The single-cell RNA-seq analysis was performed using the Tumor Immune Single-cell Hub 2 web service (http://tisch.comp-genomics.org/home/)[21] using the uniform manifold approximation and projection (UMAP) method to reduce dimensionality and visualize the clustering results[21]. The mRNA expression of different cells was also visualized using UMAP distribution figures. The R statistical software (version 3.63; http://www.r-project.org/) was used to perform statistical analysis. All statistical tests were two-sided, and a $P < 0.05$ was considered statistically significant.

Results

HRG prognostic signature construction. Using the univariate Cox regression analysis, we identified 12 prognostic hypoxia-related DEGs in the TCGA colon cancer database (Fig. 2A). A superior signature containing six prognostic HRGs was constructed based on multivariate Cox regression analyses of the TCGA (training) and GEO (testing) sets. Three protective genes were significantly associated with prognosis (Fig. 2B): hexokinase 1 (*HK1*; hazard ratio [HR] = 0.541, 95% confidence interval [CI] 0.314–0.930), aldolase B (*ALDOB*; HR = 0.792, 95% CI 0.651–0.963), and glucose-6-phosphate isomerase (*GPI*; HR = 0.633, 95% CI 0.350–1.145). In addition, three hazard genes were identified: enolase 3 (*ENO3*; HR = 2.643, 95%CI 1.362–5.129), serpin family E member 1 (*SERPINE1*; HR = 1.470, 95% CI 1.187–1.821), and transketolase-like protein 1 (*TKTL1*; HR = 1.306, 95% CI 1.059–1.621). The coefficients were also verified by multivariate Cox analysis (derived using the forward–backward method; Table 1). The CRC patients' risk scores were calculated as follows: (0.971881 × expression [*ENO3*]) + (0.385271 × expression [*SERPINE1*]) + (0.267247 × expression [*TKTL1*]) − (0.61523 × expression [*HK1*]) − (0.23354 × expression [*ALDOB*]) − (0.45796 × expression [*GPI*]).

Prognostic HRG signature validation. We evaluated each COAD case's risk score and divided patients into two risk groups (low and high) according to the median risk score. K–M curves indicated that the high-risk group had a poorer OS rate than the low-risk group in both training (TCGA cohort; Fig. 2C) and testing (GSE39582; Fig. 2D) sets. In addition, another CRC cohort (GSE103479) was used to further validate the HRG signature's prognostic value, with the high-risk group showing worse prognoses than the low-risk group (Fig. 2E). The HRG signature's area under the ROC curve (AUC) indicated good predictability at one (AUC = 0.693), three (AUC = 0.712), and five (AUC = 0.780) years (Fig. 2F). Then, univariate and multivariate Cox regression analyses were used to estimate whether the risk model could be an independent clinically prognostic factor. The results showed that age, sex, American Joint Committee on Cancer (AJCC) stage, tumor (T) stage, node (N) stage, and metastasis (M) stage were significantly correlated with the risk score (Fig. 2G).

Furthermore, the high-risk group contained a higher percentage of cases with death outcomes than the low-risk group in both TCGA (high risk: 30% with dead status, 70% with alive status; low risk: 8% with dead status, 92% with alive status) and GEO (high risk: 43% with dead status, 53% with alive status; low risk: 30% with dead status, 70% with alive status) datasets (Fig. 3A,B). Using correlation analyses, we found that the signature's six HRGs were independent in both databases (Fig. 3C,D). Moreover, the expression of key genes in both datasets was plotted in the heatmap (Fig. 3E,F).

Functional annotation. In addition, we analyzed the expression of the signature's six HRGs in different tumor stages. *ALDOB* expression was significantly lower at T4 and N2 (Fig. 4A,B). Using GSEA enrichment analysis, we identified six pathways enriched in the high-risk group, including extracellular matrix receptor interaction, focal adhesion, gap junction, glycosaminoglycan biosynthesis, and melanoma pathways. In addi-
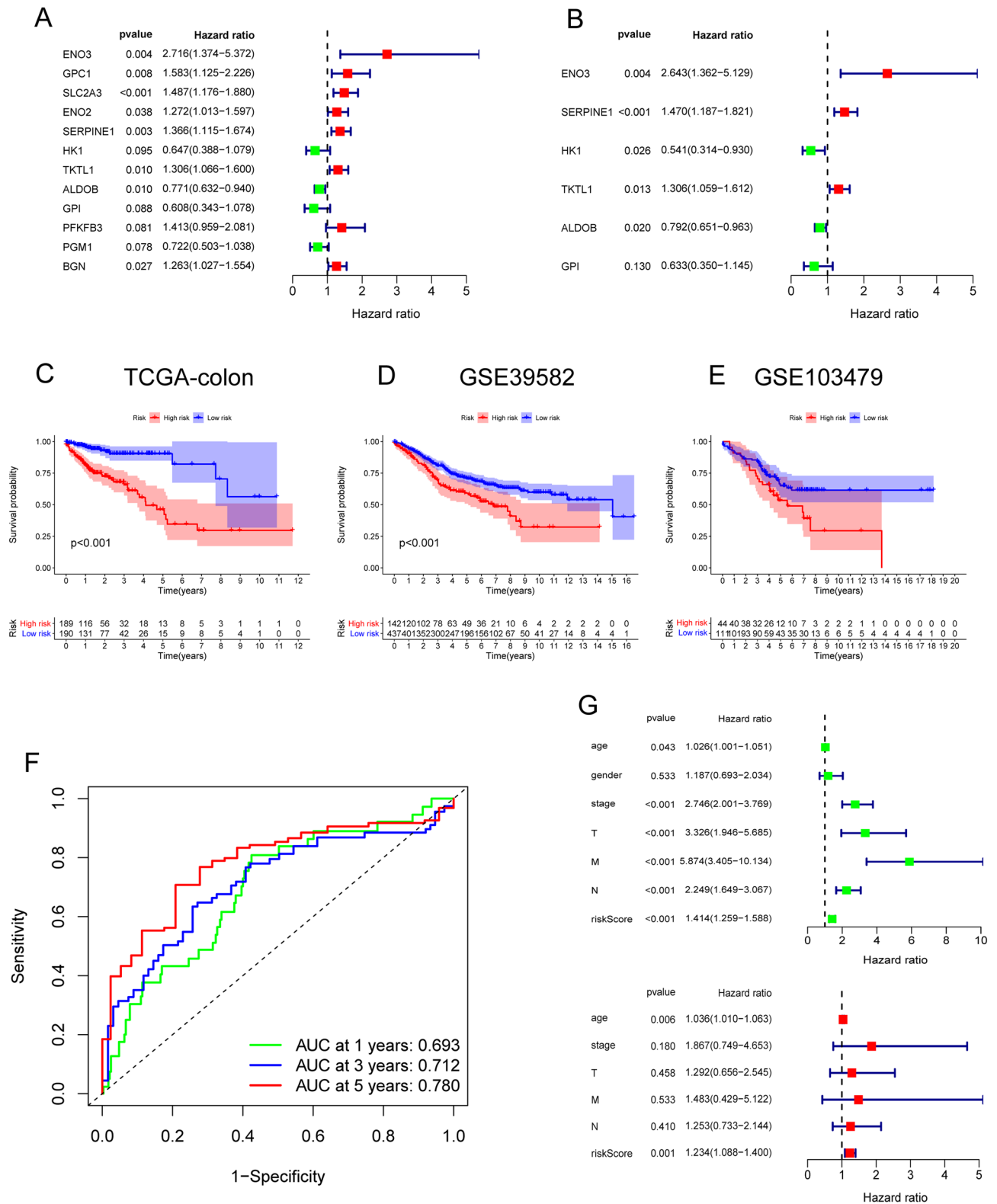
Figure 2. Construction of the prognostic HRG model. (**A**) Differentially expressed HRGs in CRC patients and healthy individuals in the TCGA. (**B**) Multivariate and univariate Cox regression analyses of the prognostic model's six HRGs. (**C**) K–M curves of OS in the TCGA dataset. (**D**) K–M curves of OS in the GEO dataset. (**E**) ROC curve for the TCGA dataset. (**F**) Multivariate and univariate Cox regression analyses of risk scores and clinicopathological characters.

| Hypoxia-related genes	Coefficients	HR	HR.95L	HR.95H	*P* value
ENO3	0.971881	2.642911	1.361901	5.128847	0.004065
SERPINE1	0.385271	1.470013	1.186897	1.820661	0.000416
HK1	− 0.61523	0.540516	0.314043	0.93031	0.026371
TKTL1	0.267247	1.306363	1.058941	1.611596	0.012613
ALDOB	− 0.23354	0.791726	0.65066	0.963376	0.019667
GPI	− 0.45796	0.632574	0.34951	1.144889	0.130292

Table 1. Coefficients of hypoxia-related genes by multivariate Cox analysis.
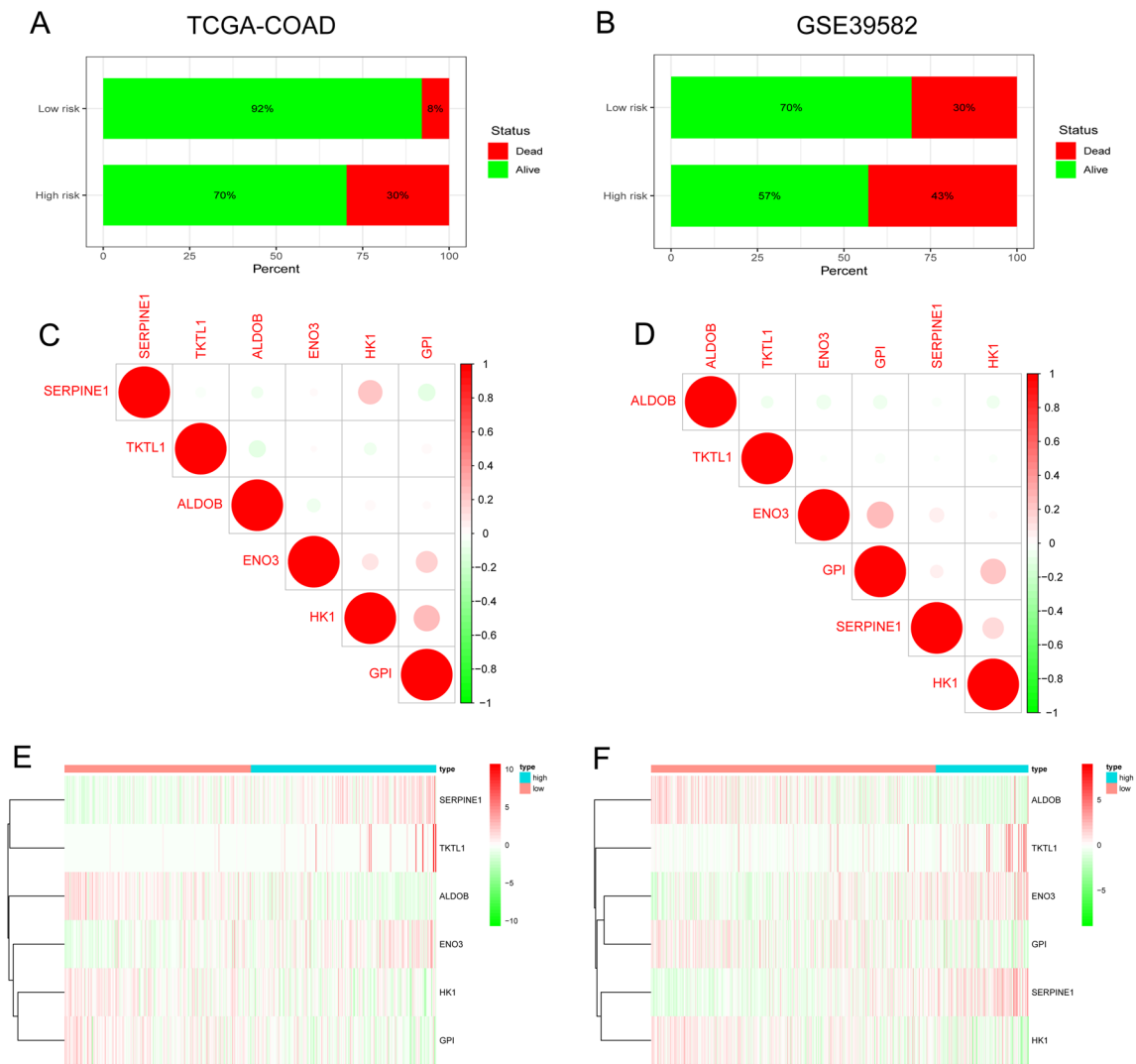
Figure 3. Prognostic HRG model verification. Percentage survival status for the low- and high-risk groups in the (**A**) TCGA and (**B**) GEO datasets. Correlation analyses of the signature's six HRGs in the (**C**) TCGA and (**D**) GEO datasets. (**E**, **F**) Heatmaps of the six HRGs in the GEO dataset.

tion, five cancer hallmarks were enriched in the high-risk group: angiogenesis, EMT, myogenesis, transforming growth factor (TGF)-β signaling, and UV response DNA pathways (Fig. 4C,D).

Immune correlation analysis. The expression of 16 TICs in the TCGA database indicated that the high-risk group had lower regulatory T cell expression and higher neutrophil expression than the low-risk group (Fig. 5A,B). Moreover, 18 immune-related genes were closely associated with risk scores in the TCGA dataset (Fig. 5C) and 23 in the GEO dataset (Fig. 5D). Additionally, we explore the underlying relationship between risk score and chemokine expression in both datasets. *CXCL9* expression was significantly increased in the high-risk
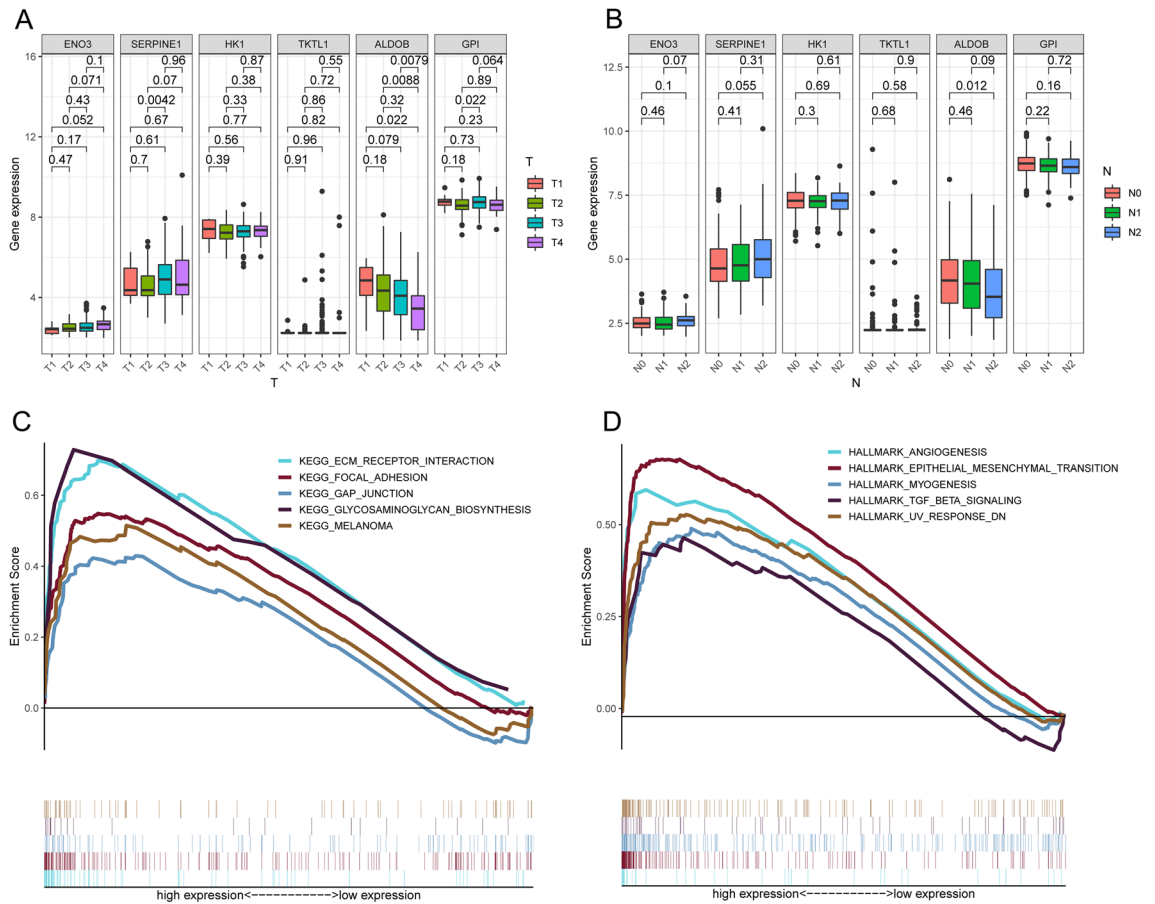
Figure 4. Function enrichment. Differential expression of the signature's six HRGs in (**A**) T and (**B**) N stage tumors. (**C**, **D**) GSEA enrichment analyses.

group in both databases. In contrast, the high-risk group had higher *CX3CL1* expression in the TCGA dataset but higher *CXCL10* expression in the GEO dataset (Fig. 5E,F).

Single-cell RNA-seq analysis of HRGs.

The different TICs were grouped by the UMAP method (Fig. 6A), and their expression of the signature's six HRGs is shown in Fig. 6B-F. *ENO3*, *SENPINE1*, and *TKTL1* expression did not vary among TICs (Fig. 6B–D). However, while *GPI* and *HK1* were universally expressed in all TICs, their levels were elevated in NK and CD8+T cells (Fig. 6E,F). However, *ALDOB* was not expressed in any TICs.

HRG expression validation in CRC tissues.

We collected ten tumor-normal tissue pairs from patients in Xijing Hospital. After extracting the tissues' mRNA, RT-PCR analysis showed *ENO3*, *SENPINE1*, and *GPI* upregulation and *HK1* and *ALDOB* downregulation in cancerous compared to adjacent noncancerous tissues (Fig. 7).

Discussion

In this study, we first studied the mRNA levels of 51 reported HRGs in CRC and normal tissues, finding 12 differentially expressed. To further assess their prognostic value, we constructed a six-gene risk signature via Cox univariate analysis, which we then validated in a GEO dataset. Functional enrichment analysis showed that the DEGs between the low- and high-risk groups were related to several pathways. We examined immune cell infiltration and immune-related genes in the low- and high-risk groups.

This study identified six prognostic HRGs: *ENO3*, *SERPINE1*, *HK1*, *TKTL1*, *ALDOB*, and *GPI*. ENO3 is a major enzyme participating in CRC glycolysis. Low *ENO3* expression has been reported to correlate significantly with prolonged OS in CRC patients. Therefore, *ENO3* might be a promising prognostic biomarker for CRC patients[22]. Besides CRC, ENO3 also functions in pancreatic ductal adenocarcinoma (PDAC). To better predict the OS of PDAC patients, a prognostic nomogram involving ENO3 might contribute to the individualized management of PDAC patients[23]. *ENO3* overexpression results directly from the loss of STK11 function. Moreover, *ENO3* knockdown had a selective anticancer effect in *STK11* mutant lung cells. Therefore, an ENO3-based therapy might be promising for patients with *STK11* mutant lung cancer[24]. ENO3 was enriched in the above HIF-1 pathway.

SERPINE1 is a serine proteinase inhibitor reported to function as a crucial extracellular matrix remodeling regulator. Compelling evidence indicates that SERPINE1 is intimately associated with poor prognoses in diverse
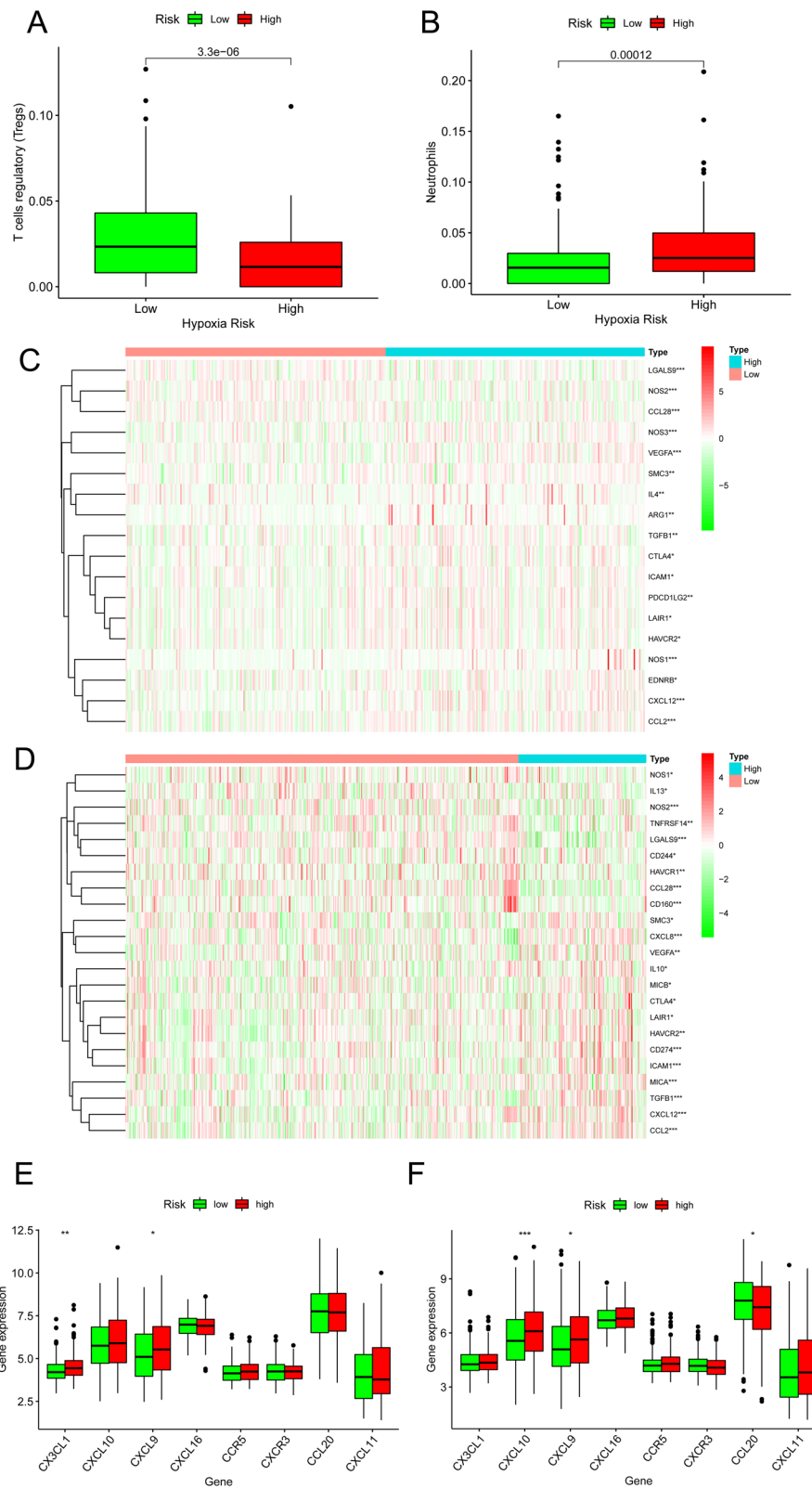
Figure 5. TICs and immune genes analysis. Expression of (**A**) regulatory T cells and (**B**) neutrophils in the TCGA dataset. (**C**, **D**) Heatmaps of immune-related gene expression. Chemokine expression in the (**E**) TCGA and (**F**) GSE39582 datasets.
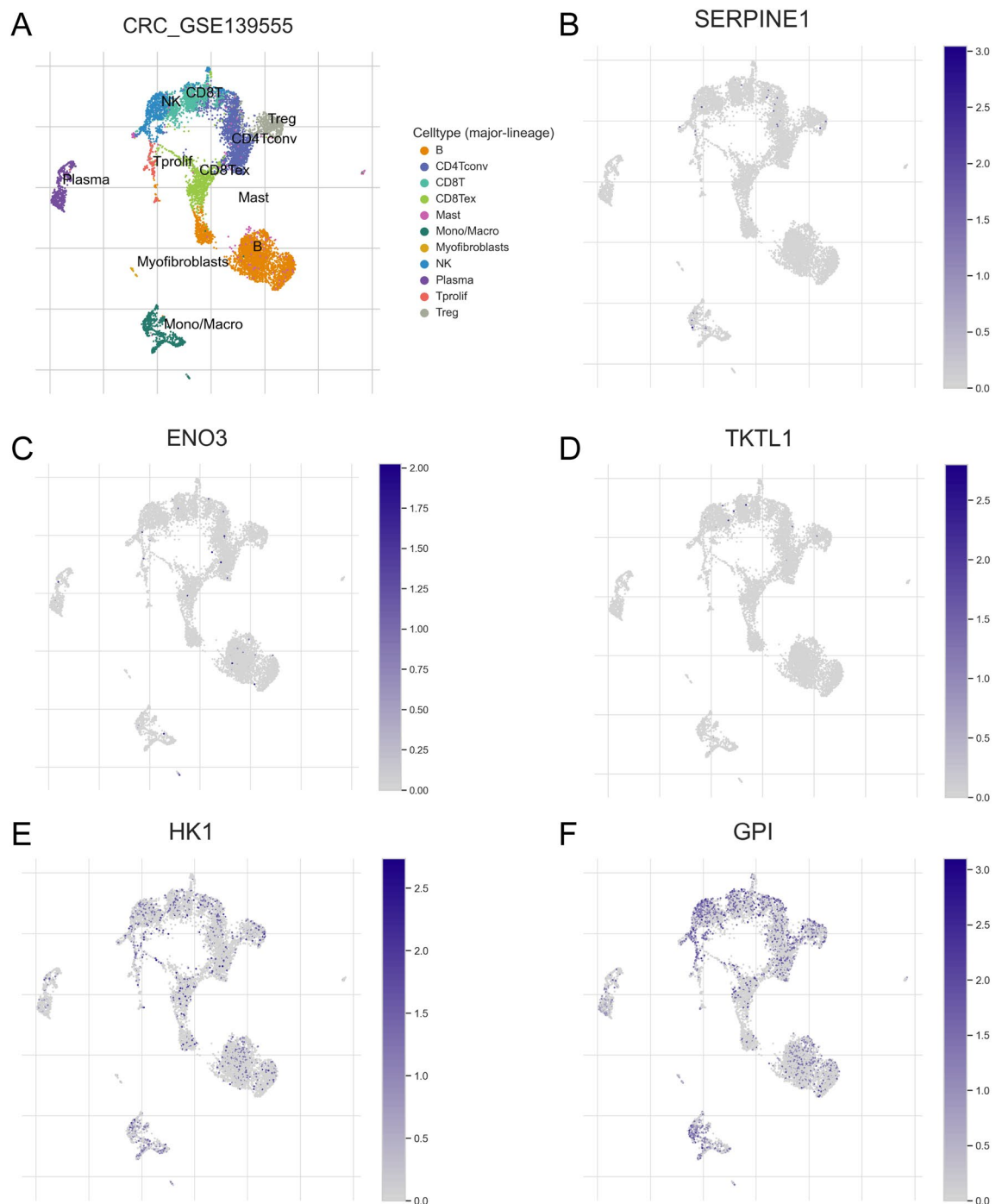
Figure 6. Single-cell RNA-seq analysis of HRGs. (**A**) The distribution of TICs by the UMAP method. The mRNA expression of (**B**) *SERPINE1*, (**C**) *ENO3*, (**D**) *TKTL1*, (**E**) *HK1*, and (**F**) *GPI* in the different TICs.

cancers. Biochemical analysis showed upregulated SERPINE1 expression in mesenchymal lung cancer. Crosstalk between TGF-β and Yes-associated protein (YAP) signaling pathways might be a precondition for this process[25]. SERPINE1 facilitates neoplastic cell proliferation, migration, and invasion by regulating EMT. High *SERPINE1* expression shortened OS in GAC patients. Therefore, SERPINE1 might function as a novel biomarker and independent prognosticator in GAC patients[26,27]. Hypoxia-mediated ROS upregulated *SERPINE1* expression in MDA-MB-468 breast cancer cells. Therefore, ROS may be an underlying therapeutic strategy for managing breast cancer metastasis[28].

HK1 participates in distinct biological processes contributing to glycolysis in cancer. HK1 might underlie oncogenes in many tumors[29]. *HK1* overexpression results in pyruvate and lactate overproduction. A detailed study into bladder cancer's metabolic phenotype might provide novel therapeutic strategies[30]. Previous studies have indicated that enhanced glycolytic activity promotes chemotherapeutic resistance in malignant cancers.
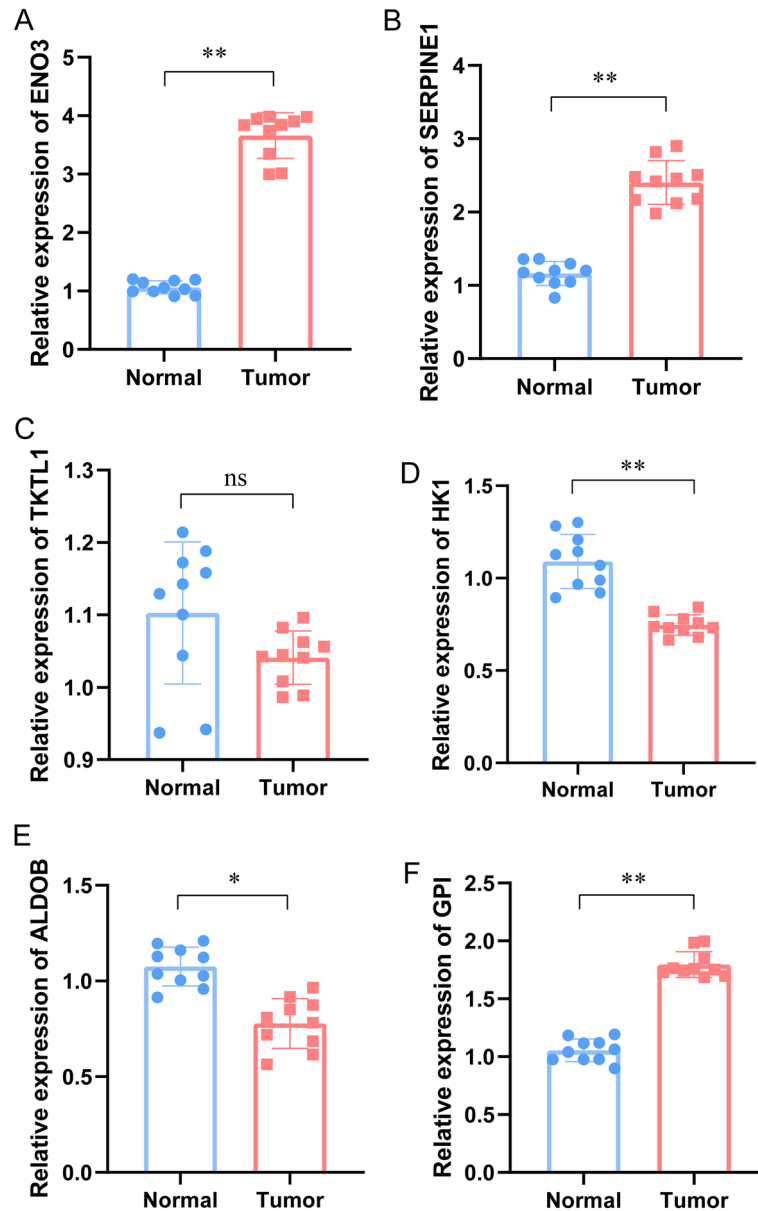
Figure 7. Expression of the signature's six HRGs in normal and CRC tissues. The mRNA expression of (**A**) *ENO3*, (**B**) *SERPINE1*, (**C**) *TKTL1*, (**D**) *HK1*, (**E**) *ALDOB*, and (**F**) *GPI*. Key: *, $P < 0.05$; **, $P < 0.01$.

However, whether glycolysis affects gemcitabine's (GEM) curative efficacy in pancreatic cancer (PC) is largely unknown. Under hypoxic conditions, *HK1* expression correlated positively with PC progression, suggesting that upregulated glycolytic activity could facilitate PC progression and increase GEM tolerance[31].

TKTL1 is an enzyme participating in cancer glycolytic metabolism. TKTL1 was a biomarker for distinguishing CRC patients who might benefit from liver resection from those who might need more aggressive perioperative chemical therapy[32]. TKTL1 participates in aerobic glycolysis in cancer, contributing to basal membrane destruction and cancer metastasis. Besides APO10, TKTL1 is another biomarker underlying epitope detection in monocytes (EDIM) blood tests. By comparing the sensitivity of the EDIM blood test with conventional tumor biomarkers, studies have shown that this novel test might aid the detection and diagnosis of cholangiocellular carcinoma, PC, and CRC[33].

The aldolase family participates in cancerous metabolism and glycolysis and has three members: ALDOA, ALDOB, and ALDOC. Abnormal aldolase levels are closely associated with several tumor types. ALDOB successfully enhanced glucose uptake and aerobic glycolysis and significantly reduced mitochondrial oxidative phosphorylation in CRC. ALDOB also altered the response of traction force in CRC, which is closely associated with cancer metastasis during hypoxia. Therefore, elevated ALDOB levels could drive hypoxia and stiff substrate to promote aerobic glycolysis in CRC, further deepening our understanding of its role in CRC advancement from a biophysical standpoint[34].

In hypoxia, GPI maintains glucose metabolism by redirecting the glucose from androgen/androgen receptor (AR)-dependent to hypoxia-induced glycolysis, attenuating chemotherapy efficacy in prostate cancer. However, GPI transcription is suppressed by AR during hypoxia. When GPI was inhibited, therapeutic resistance under hypoxia was reduced, and enzalutamide efficacy was increased[35]. Upregulated glycolysis could improve organ tolerance to hypoxia. Therefore, suppressing glycolysis-related enzyme activity and energy metabolism might be a promising treatment strategy for malignant tumors. One study showed that esculetin could disturb glucose metabolism by binding with GPI, playing an antitumor role[34].

Altogether, we have shown that hypoxia is closely prognostically associated with CRC and constructed an easy-to-use six HRG signature that is strongly associated with the tumor microenvironment in CRC. Furthermore, we used multiple datasets in order to compensate for the limitations of the individual datasets. This risk signature was an independent risk factor for predicting OS in both TCGA and GEO CRC cohorts. The COX algorithms have limited capabilities and can be supplemented with artificial intelligence algorithms. The DEGs between low and high-risk CRC groups were associated with TICs and immune-related genes. Our study provides a novel six HRG signature for predicting the overall prognosis of CRC patients and provides a robust basis for future studies on the relationship between HRGs and immunity.

Data availability

References

1. Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**(3), 209–249. https://doi.org/10.3322/caac.21660 (2021).
2. Brody, H. Colorectal cancer. *Nature* **521**(7551), S1. https://doi.org/10.1038/521S1a (2015).
3. Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M. & Wallace, M. B. Colorectal cancer. *Lancet* **394**(10207), 1467–1480. https://doi.org/10.1016/s0140-6736(19)32319-0 (2019).
4. Shao, C. *et al.* Role of hypoxia-induced exosomes in tumor biology. *Mol. Cancer* **17**(1), 120. https://doi.org/10.1186/s12943-018-0869-y (2018).
5. Jing, X. *et al.* Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol. Cancer* **18**(1), 157. https://doi.org/10.1186/s12943-019-1089-9 (2019).
6. Peng, S. *et al.* The VHL/HIF axis in the development and treatment of pheochromocytoma/paraganglioma. *Front. Endocrinol. (Lausanne)* **11**, 586857. https://doi.org/10.3389/fendo.2020.586857 (2020).
7. Kung-Chun Chiu, D. *et al.* Hypoxia regulates the mitochondrial activity of hepatocellular carcinoma cells through HIF/HEY1/PINK1 pathway. *Cell Death Dis.* **10**(12), 934. https://doi.org/10.1038/s41419-019-2155-3 (2019).
8. Zhou, Y. *et al.* Meloxicam, a selective COX-2 inhibitor, mediates hypoxia-inducible factor- (HIF-) 1α signaling in hepatocellular carcinoma. *Oxid. Med. Cell Longev.* **2020**, 7079308. https://doi.org/10.1155/2020/7079308 (2020).
9. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102 (2005).
10. Zhu, J. *et al.* CXCL5/CXCL8 is a promising potential prognostic and tumor microenvironment-related cluster in hepatocellular carcinoma. *J. Gastrointest. Oncol.* **11**(6), 1364–1380. https://doi.org/10.21037/jgo-20-556 (2020).
11. Zhu, J. *et al.* Comprehensive analysis of the relationship between competitive endogenous RNA (ceRNA) networks and tumor infiltrating-cells in hepatocellular carcinoma. *J. Gastrointest. Oncol.* **11**(6), 1381–1398. https://doi.org/10.21037/jgo-20-555 (2020).
12. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
13. Zhu, J. *et al.* Identification of five glycolysis-related gene signature and risk score model for colorectal cancer. *Front. Oncol.* **11**, 588811 (2021).
14. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**(5), 453–457. https://doi.org/10.1038/nmeth.3337 (2015).
15. Pan, X., Wu, H., Chen, G. & Li, W. Prognostic value of enolase gene family in colon cancer. *Med. Sci. Monit.* **26**, e922980. https://doi.org/10.12659/msm.922980 (2020).
16. Tan, Z. *et al.* The value of a metabolic reprogramming-related gene signature for pancreatic adenocarcinoma prognosis prediction. *Aging (Albany NY)* **12**(23), 24228–24241. https://doi.org/10.18632/aging.104134 (2020).
17. Park, C. *et al.* Overexpression and selective anticancer efficacy of ENO3 in STK11 mutant lung cancers. *Mol. Cells* **42**(11), 804–809. https://doi.org/10.14348/molcells.2019.0099 (2019).
18. Carreras, J., Hamoudi, R. & Nakamura, N. Artificial intelligence analysis of gene expression data predicted the prognosis of patients with diffuse large B-cell lymphoma. *Tokai J. Exp. Clin. Med.* **45**(1), 37–48 (2020).
19. Zhang, B. *et al.* Comparative transcriptomic and proteomic analyses provide insights into the key genes involved in high-altitude adaptation in the Tibetan pig. *Sci. Rep.* **7**(1), 3654. https://doi.org/10.1038/s41598-017-03976-3 (2017).
20. Kong, H. J. *et al.* Crosstalk between YAP and TGFβ regulates SERPINE1 expression in mesenchymal lung cancer cells. *Int. J. Oncol.* **58**(1), 111–121. https://doi.org/10.3892/ijo.2020.5153 (2021).
21. Yang, J. D., Ma, L. & Zhu, Z. SERPINE1 as a cancer-promoting gene in gastric adenocarcinoma: Facilitates tumour cell proliferation, migration, and invasion by regulating EMT. *J. Chemother.* **31**(7–8), 408–418. https://doi.org/10.1080/1120009x.2019.1687996 (2019).
22. Li, L. *et al.* FN1, SPARC, and SERPINE1 are highly expressed and significantly related to a poor prognosis of gastric adenocarcinoma revealed by microarray and bioinformatics. *Sci. Rep.* **9**(1), 7827. https://doi.org/10.1038/s41598-019-43424-x (2019).
23. Liao, P. *et al.* Genome-scale analysis identifies SERPINE1 and SPARC as diagnostic and prognostic biomarkers in gastric cancer. *Oncol. Targets Ther.* **11**, 6969–6980. https://doi.org/10.2147/ott.S173934 (2018).
24. Azimi, I., Petersen, R. M., Thompson, E. W., Roberts-Thomson, S. J. & Monteith, G. R. Hypoxia-induced reactive oxygen species mediate N-cadherin and SERPINE1 expression, EGFR signalling and motility in MDA-MB-468 breast cancer cells. *Sci. Rep.* **7**(1), 15140. https://doi.org/10.1038/s41598-017-15474-7 (2017).
25. Jiang, M. *et al.* A pan-cancer analysis of molecular characteristics and oncogenic role of hexokinase family genes in human tumors. *Life Sci.* **264**, 118669. https://doi.org/10.1016/j.lfs.2020.118669 (2021).

26. Massari, F. *et al.* Metabolic phenotype of bladder cancer. *Cancer Treat. Rev.* **45**, 46–57. https://doi.org/10.1016/j.ctrv.2016.03.005 (2016).
27. Dai, S. *et al.* Glycolysis promotes the progression of pancreatic cancer and reduces cancer cell sensitivity to gemcitabine. *Biomed. Pharmacother.* **121**, 109521. https://doi.org/10.1016/j.biopha.2019.109521 (2020).
28. Peltonen, R. *et al.* High TKTL1 expression as a sign of poor prognosis in colorectal cancer with synchronous rather than metachronous liver metastases. *Cancer Biol. Ther.* **21**(9), 826–831. https://doi.org/10.1080/15384047.2020.1803008 (2020).
29. Saman, S. *et al.* Biomarkers Apo10 and TKTL1: Epitope-detection in monocytes (EDIM) as a new diagnostic approach for cholangiocellular, pancreatic and colorectal carcinoma. *Cancer Biomark* **27**(1), 129–137. https://doi.org/10.3233/cbm-190414 (2020).
30. Ma, L. *et al.* CircDUSP16 contributes to cell development in esophageal squamous cell carcinoma by regulating miR-497-5p/TKTL1 Axis. *J. Surg. Res.* **260**, 64–75. https://doi.org/10.1016/j.jss.2020.11.052 (2021).
31. Chang, Y. C., Yang, Y. C., Tien, C. P., Yang, C. J. & Hsiao, M. Roles of aldolase family genes in human cancers and diseases. *Trends Endocrinol. Metab.* **29**(8), 549–559. https://doi.org/10.1016/j.tem.2018.05.003 (2018).
32. Huang, H. C. *et al.* Aldolase triggers metabolic reprogramming in colorectal cancer in hypoxia and stiff desmoplastic microenvironments. *Colloids Surf. B Biointerfaces* **190**, 110969. https://doi.org/10.1016/j.colsurfb.2020.110969 (2020).
33. Geng, H. *et al.* Publisher correction: Interplay between hypoxia and androgen controls a metabolic switch conferring resistance to androgen/AR-targeted therapy. *Nat. Commun.* **10**(1), 164. https://doi.org/10.1038/s41467-018-07872-w (2019).
34. Wu, S. T. *et al.* Esculetin inhibits cancer cell glycolysis by binding tumor PGK2, GPD2, and GPI. *Front. Pharmacol.* **11**, 379. https://doi.org/10.3389/fphar.2020.00379 (2020).
35. Geng, H. *et al.* Publisher Correction: Interplay between hypoxia and androgen controls a metabolic switch conferring resistance to androgen/AR-targeted therapy. *Nat. Commun.* **10**, 164 (2019).

Author contributions

Q.Y. wrote the main manuscript text. J.X. conducted data statistics and analysis. W.K. and C.R. prepared figures. S.L. and L.J. proposed the research design. L.Y. and L.J. reviewed the knowledge content of the article. All authors reviewed the manuscript. These authors contributed equally: Y.Q. and X.J.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-29718-2.

Correspondence and requests for materials should be addressed to L.S. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.