

Identification and Validation of a Hypoxia-related lncRNA Signature for Prognostic Prediction in Endometrial Cancer

Hongyang Liu

Zhengzhou University

Junhu Wan

Zhengzhou University

Quanling Feng

Zhengzhou University

Jingyu Li

zhengzhou people's hospital

Jun Liu

Shangqiu Medical College

Shihong Cui (✉ cuishzzuedu@163.com)

Zhengzhou University <https://orcid.org/0000-0002-8996-7255>

Research

Keywords: Endometrial cancer, SOS1-IT1, hypoxia, lncRNA, prognostic prediction

Posted Date: June 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-633056/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Endometrial cancer (EC) is one of the most common types of gynecological cancer. Hypoxia is an important clinical feature and regulates various tumor processes. However, the prognostic value of hypoxia-related lncRNA in EC remains to be further elucidated. Here, we aimed to characterize the molecular features of EC by the development of a classification system based on the expression profile of hypoxia-related lncRNA.

Methods: Univariate Cox regression analysis was used to identify hypoxia-related lncRNAs associated with overall survival. The least absolute shrinkage and selection operator (LASSO) Cox regression model was utilized to construct gene signature. Multivariate Cox regression analysis and receiver operating characteristic (ROC) curve analysis were also performed. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEEG) pathway, and Gene Set Enrichment Analysis (GESA) were used to identify hypoxia-related lncRNA pathway. Western blot and real-time PCR were used to detect target gene expression. The cell proliferation was determined by using WST-1 assay.

Results: Based on univariate Cox regression analysis, we identified 17 hypoxia-related lncRNAs significantly associated with overall survival. Next, the least absolute shrinkage and selection operator (LASSO) Cox regression model was utilized to construct a multigene signature in the TCGA EC cohort. The risk score was confirmed as an independent predictor for overall survival in multivariate Cox regression analysis and receiver operating characteristic (ROC) curve analysis. Besides, the survival time of EC patients in different risk group was significantly correlated to clinicopathologic factors, such as age, stage and grade. Furthermore, hypoxia-related lncRNA associated with the high-risk group were involved in various aspects of the malignant progression of EC via Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEEG) pathway, and Gene Set Enrichment Analysis (GESA). Besides, using CIBERSORT analysis, we found a different immune cell environment characterization of EC between different cluster and risk group. Moreover, the risk score was closely correlated to immunotherapy response, microsatellite instability and tumor mutation burden (TMB). Finally, we select one hypoxia-related lncRNA SOS1-IT1 to validate its role in hypoxia and EC progression. Interestingly, we found SOS1-IT1 was overexpressed in tumor tissues, and closely correlated with clinicopathological parameters of EC. The expression level of SOS1-IT1 was significantly increased under hypoxia condition. Additionally, the important hypoxia regulatory factor HIF-1 α can directly bind SOS1-IT1 promoter region, and affect its expression level.

Conclusions: In summary, this study established a new EC classification based on the hypoxia-related lncRNA signature, thereby provide a novel sight to understand the potential mechanism of human EC development.

Background

Endometrial cancer (EC) is one of the most common gynecological cancer in developed countries[1, 2]. EC has been considered as a global threat to women's health and well-being. Most EC patients are diagnosed at an early stage with a favorable prognosis. However, 20-30% of EC patients diagnosed with a high-risk of recurrence and poor prognosis. EC patients often showed different prognoses and treatment responses, even if with the same degree of progression[3-5]. Therefore, effective EC biomarkers must be found to assess prognosis and identify EC patients with high-risk.

Hypoxia is a common phenomenon in solid tumors. Hypoxia is harmful to cancer cells, but it can promote cancer cells to adapt, thus promoting malignant progression[6]. In the hypoxic microenvironment, tumor cells exhibit activating downstream genes to facilitate tumor growth, with strong aggressiveness and metastasis ability[7, 8]. In response to hypoxia, the major component of hypoxia signaling pathways is hypoxia inducible factors-1 (HIF-1), which is a heterodimer composed of α and β subunits[9, 10]. HIF-1 α mainly interact with HIF-1 β subunit to form a stable HIF-1 complex in the nucleus. Under hypoxic conditions, the HIF-1 complex binds to the target gene's promoter to induce their transcription. Most of the HIF-1 complex dependent genes were associated with proliferation, epithelial to mesenchymal transition, angiogenesis, and metastasis[11-13]. However, due to the lack of effective biomarkers, how hypoxia leads to tumor progression remain to be further elucidated.

Long non coding RNAs (lncRNAs) are a large class of heterogeneous transcripts with a length of more than 200 nucleotides, and their protein coding potential is limited. lncRNAs are widely transcribed in eukaryotic genomes, suggesting their important regulatory role in complex organisms[14, 15]. Although only a small number of functional lncRNAs have been well characterized, more and more evidence showed that lncRNAs play a key role in controlling a large number of cancer-related cell processes, such as proliferation, migration, invasion, autophagy and stemness[16-19]. A specific group of lncRNAs is regulated by tumor microenvironment, such as hypoxia. The hypoxia responsive lncRNAs, such as NORAD, RAB11B-AS1 and lncHIFCAR, may be the basis of cancer cell survival and disease progression[20-22]. MALAT1 was the first reported lncRNA mediating hypoxia-induced pro-survival autophagy in endometriosis[23]. However, there is few reports about the relationship between hypoxia related lncRNA and EC in recent years. Whether lncRNAs are involved in the response to hypoxia in EC, and the prognostic value of hypoxia-related lncRNA in EC remains to be further elucidated.

In this study, we first downloaded the RNA expression profiles of hypoxia-related lncRNA and corresponding clinical data of EC patients from public databases. The different lncRNA expression patterns were detected among EC cases, to identify the candidate lncRNA biomarkers based on The Cancer Genome Atlas data (TCGA). Besides, we found that the hypoxia-related lncRNA could classify the EC patients with significantly different clinical and molecular characteristics. Finally, we focused on lncRNA SOS1-IT1, and found it was significantly upregulated in EC cells under hypoxic conditions. We evaluated its biological role and clinical significance in EC progression and revealed SOS1-IT1 is hypoxia-inducible and directly transactivated by HIF-1 α .

Materials And Methods

Data collection

The TCGA RNA sequencing dataset and corresponding clinicopathological characteristics information, such as age, grade, stage, radiation therapy, surgical approach and survival information were downloaded from TCGA database (<http://www.cancergenome.nih.gov/>).

Identification of differentially expressed hypoxia-related lncRNA

Based on the Spearman correlation analysis, lncRNAs related to the hypoxia were identified (Filter: $|r| > 0.5$ and $P < 0.001$). The prognosis-related lncRNAs were screened by K-M survival analyses ($P < 0.01$). Differentially expressed hypoxia-related lncRNA were identified in EC from the TCGA datasets by using the EdgeR package in R statistical software. The significance criteria for determining differentially expressed lncRNA were set as adjusted P value < 0.05 . Heatmaps were plotted by the “pheatmap” package.

Comprehensive analysis of interaction network

According to the correlation efficiency and probability cut-off value (Filter: $|r| > 0.5$ and $P < 0.001$), 677 lncRNAs were screened out as hypoxia-related lncRNAs. Then we explored the interaction network between hypoxia and hypoxia-related lncRNAs, by using the Cytoscape software.

Generation of hypoxia-related lncRNA signature

The expression data of hypoxia-related lncRNA in 550 EC and 35 normal tissues were analyzed with the Limma package and visualized as a heatmap. We carried out consensus clustering with the R package “ConsensusClusterPlus.” Then we used “Ggplot2” and “Limma” package for PCA analysis. The “Glmnet” and “Survival” packages were used to perform LASSO regression analysis[24]. We also performed the univariate and multivariate Cox hazard analysis of clinicopathological characteristics information by “survival” package.

Construction of hypoxia-related lncRNA prognostic model

Prognosis-related lncRNA were constructed using multivariate cox regression. The risk score of each patient was calculated according to the predictive signature model. Using the median risk score as the cutoff point, EC patients were divided into the high-risk group and low-risk group. The prediction efficiency of the risk signature model was evaluated by the receiver operating characteristic (ROC) curve using R package survival ROC.

Functional enrichment analysis of the hypoxia-related lncRNA.

Functional enrichment analysis of the hypoxia-related lncRNA was conducted using DAVID, including biological functions, cellular components, and molecular functions. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was searched for significant pathways. We also performed Gene set enrichment analysis (GSEA) to identify the significantly alerted KEGG pathways between the high-risk and

low-risk groups. The Java GSEA program were used and the gene set from Molecular Signatures Database was selected as the reference gene set. Biological processes with normalized p value < 0.05 and false discovery rate (FDR) q value < 0.05 were considered statistically significant.

Evaluation of tumor-infiltrating immune cells and the immune infiltration level

The subpopulation of 22 immune cells in each tumor sample was explored by CIBERSORT algorithm in different clusters. The CIBERSORT results of samples with p < 0.05 showed that the inferred fractions of immune cell populations were accurate and could be further analyzed. Furthermore, we analyzed the correlation between tumor immune cell infiltrating and prognostic risk signature.

Tumor mutation burden assessment

The tumor mutation burden (TMB) for each organization was detected by the VarScan method, as calculated using the R package "maftools". According to the cutoff value of median TMB data, EC Patients were divided into high TMB group and low TMB group. We also analyzed the correlation between TMB and prognostic risk signature.

Cell culture

Human endometrial cancer cell lines Ishikawa was purchased from ATCC (American Type Culture Collection). The cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and cultured at 37°C with 5% CO₂.

Total RNA extraction and Real-time PCR

The total RNA was extracted by TRIZOL reagent (Invitrogen) according to the manufacturer's instruction. The cDNA was obtained from the purified RNA using a PrimeScript RT Reagent Kit (Takara). Real-time PCR was performed using the SYBR Premix Ex Taq (Takara) following the manufacturer's instructions. The results were normalized to β -actin gene.

Cell proliferation assays

The cell proliferation assays were determined using a commercial WST-1 assay kit according to the manufacturer's instructions as described before[25-27]. First, 15 μ l WST-1 reagent was added to the cells in the 96-well plates, and incubated for 2 h. Later, the OD value was measured in a microplate reader at 450 nm.

Dual-luciferase reporter assays

For the luciferase reporter experiment, the SOS1-IT1 wild type (WT) or mutant vectors were used. The Ishikawa cells were seeded in six-well plates and transfected with relative plasmids or siRNA for 48 h. The Dual-Luciferase Reporter Assay System was used to assess the activities of firefly luciferase and Renilla

luciferase sequentially. Then the relative luciferase activities were calculated, and control cells were used for normalization.

Statistical analysis

Differences between survival curves were generated by the Kaplan-Meier method and compared by log-rank tests. The multivariate analysis was performed using the cox proportional hazard model. For comparisons of two groups, a t-test was used. R studio package was used for all statistical analysis. All statistical tests were only considered statistically significant when $p < 0.05$ was achieved.

Results

Identification hypoxia-related lncRNA in EC

To better understand the important role of hypoxia-related lncRNA in oncogenesis and progression, we first downloaded from the hallmark gene sets of hypoxia including 200 genes from Molecular Signatures Database. RNA-seq data from 550 tumor tissue samples and 35 normal samples were downloaded from TCGA. According to the correlation efficiency and probability cut-off value (Filter: $|r| > 0.5$ and $P < 0.001$), 677 lncRNAs were screened out as hypoxia-related lncRNA. In order to investigate the prognostic value of these hypoxia-related lncRNA in EC, univariate Cox regression analysis was performed based on the expression levels of these lncRNA in TCGA database. As a result, we found that 17 out of the 677 lncRNAs were significantly associated with overall survival ($p < 0.05$) (Figure 1A). The expression values of 17 hypoxia-related lncRNAs were extracted from EC patients. The expression profile of 17 prognostic associated hypoxia-related lncRNA was showed in the heatmap and box plot (Figure 1B-C). As shown in Figure 1C, 17 hypoxia-related lncRNAs were significantly abnormally expressed in EC tissues samples. Compared to normal tissue samples, lncRNAs AL021578.1, AP000593.3, LINC01843, SOS1-IT1 and NRAV were up-regulated, while AC079921.1, AC119674.1, AC114947.2, AC011466.1, CFAP58-DT, LINC01614, AC244517.7, TPM1-AS, AL035530.2, AC007786.2, AP003096.1 and AC020910.5 were down-regulated in EC tumor tissues. In addition, based on the co-expression relationship, we also built an interaction network based on hypoxia-related lncRNAs and the corresponding hypoxia associated genes (Figure 1D).

Classification of EC based on hypoxia-related lncRNA

In order to analysis the consensus cluster of hypoxia-related lncRNA in EC, we used the common clusterplus package to identify the different groups of hypoxia-related lncRNA based on their co-expression patterns in EC tissues from the TCGA database. Due to the grouping was suboptimal when they were divided into more than 2 clusters, we divided the hypoxia-related lncRNA into two groups based on their expression indices using $k = 2$ as the optimal value (Figure 2A-D). Next, we analyzed the relationship between these two clusters and the clinicopathological characteristics of EC patients. We found that consensus clustering could make significant differences in the clinical and molecular characteristics of the two EC clusters (Figure 2E). Cluster 1 patients were strikingly correlated with stage,

grade, fustat and disease type by Chi-square test. Besides, compared with patients in cluster 2, EC patients in cluster 1 showed a shorter overall survival time (Figure 2F).

Prognostic value of hypoxia-related lncRNA and construction of a risk signature predicting prognosis

To identify the most powerful prognostic hypoxia-related lncRNAs, the last absolute shrinkage and selection operator (LASSO) Cox regression analysis to the 17 prognosis-related lncRNAs was conducted. We constructed prognostic models using the multivariate Cox proportional hazards regression analysis and the coefficient of each independent prognostic gene were calculated. The risk score was estimated based on the coefficients from the LASSO results. According to the median risk score, EC patients were assigned into low-risk and high-risk groups. The distribution of the hypoxia-related risk signature in the TCGA dataset and survival status of EC patients in different groups were shown in Figure 3A-B. After adjusting for clinicopathological features such as age, grade, stage, radiation therapy, surgical approach, and disease type, we found that age, grade, stage, disease type and risk score were correlated with the OS of EC patients in univariate analysis, while multivariate COX regression analysis showed that age, grade, stage, radiation therapy, surgical approach and risk score were independent risk factors for the prognosis of EC patients (Figure 3C-D).

The heatmap of these most significantly six lncRNAs were shown in Figure 3E, and we observed a strong correlation between the risk score and the clinicopathological characteristics such as stage, age and grade. The result of survival analysis showed that the high-risk group had significantly shorter survival time compared to low-risk group (Figure 3F). The AUC value of the ROC curve in risk score for 5-year survival is 0.691, which is obviously higher than that of ROC in grade (0.662), radiation therapy (0.371), surgical approach (0.461) and disease type (0.569), but lower than stage (0.702) (Figure 3G). In addition, the calibration plot revealed ideal agreements of these genes between the actual observations and 1-, 3- and 5-year OS predicted by the nomogram (Figure 3H). These results indicated that the prognostic index based on hypoxia-related lncRNA has the potential to predict the survival of EC patients.

Relationships between the risk score and clinicopathologic factors

Next, we built a complete prognostic model based on the entire set. The stratification analysis was done according to the age, grade, stage, radiation therapy and surgical approach. EC patients were stratified into age ≤ 60 and > 60 subgroup, grade G1-G2 and G3-G4 subgroup, stage I-II and III-IV subgroup, surgical approach minimally invasive and open subgroup, radiation therapy No and Yes subgroup. For the EC patients in age > 60 subgroup, grade G3-G4 subgroup, and stage III-IV subgroup, the survival time of patients was significantly shorter than that of patients in another group (Figure 4A), and the average risk score was much higher (Figure 4D-E). However, there is no difference between surgical approach and radiation therapy subgroup. Besides, we found the survival time of patients in high-risk group was significantly shorter in age > 60 subgroup, grade G3-G4 subgroup, stage III-IV subgroup, surgical approach minimally invasive subgroup, and radiation therapy No subgroup (Figure 4B-C).

Biological characteristics and pathway analysis of hypoxia-related risky lncRNA

The EC patients in TCGA were divided into high-risk and low-risk groups based on the median risk score. Then, the lncRNA significantly upregulated (fold change > 1 and $p < 0.05$) or downregulated (fold change < -1 and $p < 0.05$) were selected for Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and Gene Set Enrichment Analysis (GSEA). To elucidate the biological functions and pathways that were associated with the risk score, the hypoxia-related lncRNA between the high-risk and low-risk groups were used to perform GO enrichment and KEGG pathway analyses (Figure 5A-D). As expected, hypoxia-related lncRNAs were enriched in several cancer-related biological pathways, such as membrane potential regulation, cell-cell adhesion, and protein digestion signaling pathway, et al. The GSEA analysis results showed that different risk group was involved in multiple significantly enriched pathways, including cell cycle, wnt signaling, spliceosome, et al (Figure 5E). These results showed the two-risk group identified based on the six hypoxia-related risky lncRNAs were closely associated with the malignancy of EC.

The correlation of risk score with immune cell environment characterization in EC

To clarify the potential role of the two clusters (Cluster 1 and 2) in immune cell environment characterization of EC, we investigated the expression levels of the 22 immune cell types infiltration between the two clusters and immune signature by CIBERSORT analysis. As shown in Figure 6A, Cluster 2 patients exhibited higher levels of proportions in CD4 memory resting T cells, activated NK cells and M0 Macrophages. While a higher proportion of CD8 T cells, follicular helper T cells, M1 Macrophages were enriched in Cluster 1 patients. After further uncovering the role of risk score in immune cell environment characterization of EC, we found the risk score was positively correlated with activated dendritic cells, naive B cells, and gamma delta T cell, which meant that hypoxia-related lncRNAs in the signature might influence the tumor immune microenvironment by promoting activated dendritic cells, naive B cells, and gamma delta T cell to infiltrate into the tumor tissue. However, the neutrophils, resting dendritic cells, regulatory T cell (Tregs), CD8 T cells, and activated NK cells were negatively correlated with risk score (Figure 6B-I).

The risk score of hypoxia-related lncRNA was associated with immunotherapy, microsatellite instability and tumor mutation burden in EC

The combination of anti-CTLA-4 and anti-PD-1 treatment could increase the proportion of activated CD8+ cells and natural killer cells in the tumor microenvironment, and decrease the proportion of inhibitory immune cells, resulting in changes in the immune landscape, which achieved therapeutic effect in the mouse model and prolonged the tumor free survival[28]. After analyzing the difference in response to immunotherapy between different risk score group. We found that low-risk group tended to respond effectively to immunotherapy such as anti-PD-1 and anti-CTLA-4 therapy (Figure 7A-D). High microsatellite instability (MSI-H) is associated with the response to immunotherapy treatment. Interestingly, we found that the hypoxia-related lncRNA risk score of EC patients with MSI-H was lower than that of EC patients with low MSI or microsatellite stability (MSS) (Figure 7E-F). Tumor genome mutation leads to the production of new antigens that is beneficial to immunotherapy. Tumor mutation

burden (TMB) is an important biomarker, which can be used to predict the response of a variety of tumors to PD-1/PD-L1 targeted immunotherapy. The box plot showed that there was a difference in mutation frequency between different risk score groups (Figure 7G). We observed that the risk score was negatively correlated with TMB in correlation analysis ($R=-0.23$, $P=1.2e-07$) (Figure 7H). The EC patients were classified into the high and low TMB group. Kaplan-Meier analysis showed High TMB significantly correlated with better prognosis (Figure 7I). Then we evaluated the synergistic effect of risk score on prognosis stratification. The survival analysis showed that there was significant difference in survival rate according to risk score between different TMB subgroups (Figure 7J).

SOS1-IT1 is clinically relevant in EC and promotes EC cell growth

SOS1-IT1 was the most correlated prognostic lncRNAs in this model. Therefore, we will further evaluate the role of SOS1-IT1 in EC to verify the hypoxia-related lncRNA model. To investigate the clinical significance of SOS1-IT1 in EC, we analyzed its expression and clinical relevance in TCGA EC database. As shown in Figure 8A-B, SOS1-IT1 was overexpressed in tumor tissues. Analysis from clinical investigations suggested that the aberrant level of SOS1-IT1 was closely correlated with clinicopathological parameters of EC, including the age, disease type, fustat and grade (Figure 8C-F). Kaplan-Meier analysis revealed that high expression of SOS1-IT1 was significantly associated with a poor prognosis in EC patients (Figure 8G). To directly investigate the biological functions of SOS1-IT1 on EC cells, SOS1-IT1 was knocked down or overexpressed respectively in EC cell line Ishikawa cells. The real-time PCR analysis results indicated that the expression level of SOS1-IT1 was markedly decreased or increased in knocking down or overexpressing cells respectively (Figure 8H-I). The downregulation of SOS1-IT1 significantly inhibited EC cell growth. In contrast, overexpressing SOS1-IT1 yielded the opposite results (Figure 8J-K). These results indicate that SOS1-IT1 may lead to increased EC aggressiveness.

SOS1-IT1 is upregulated under hypoxia and directly transactivated by HIF-1 α

To further confirm whether SOS1-IT1 was a functional effector of hypoxia in EC progression, we treated Ishikawa cells with hypoxia or its chemical inducer CoCl_2 for 24 h, and found the expression level of SOS1-IT1 was significantly increased (Figure 9A). Besides, the expression level of SOS1-IT1 was significantly decreased after knocking down HIF-1 α , which is the main signaling pathway response component to hypoxia, both in normoxia and hypoxia condition (Figure 9 B-C). After detecting of the upstream region (~1KB upstream) of SOS1-IT1 gene by promoter sequence analysis, we found a putative HIF-1 α binding site in the promoter region of SOS1-IT1 gene (Figure 9D). We generated the mutant binding site of the reporter constructs, following with luciferase assays after transfection of different reporter constructs in EC cells. Under normoxia conditions, there was no much difference of luciferase activity in SOS1-IT1 wild type (WT) or mutant, as well as empty vector group. However, under hypoxic conditions, there was a nearly six-fold induction of luciferase activity in SOS1-IT1 WT construct. Mutation in SOS1-IT1 promoter decreased the luciferase activity to nearly basal levels caused by hypoxia (Figure 9E). Besides, the Chromatin immunoprecipitation (ChIP) assays confirmed that HIF-1 α directly bind to the promoter region promoter of SOS1-IT1 (Figure 9F). Next, we knocked down HIF-1 α , and found HIF-1 α

suppression remarkably repressed the luciferase density in cells with WT promoter, but not in mutant group (Figure 9G). Moreover, SOS1-IT1 silencing partially reduced EC cell proliferation ability by hypoxia condition or HIF-1 α overexpression (Figure 9H-J). Totally, these results intensively indicated that SOS1-IT1 was upregulated under hypoxia and a direct transcriptional target of HIF-1 α .

Discussion

Endometrial cancer is the sixth most common neoplasm in females, with rapidly increasing in the worldwide and causes ~74,000 deaths per year[1, 3]. The TCGA networking group initially identified four molecular subtypes with different prognosis based on genome somatic copy number changes, microsatellite instability and tumor mutation load[29, 30]. There are more and more molecular targeted therapies and diagnostic methods for EC[31]. However, the clinical effect of EC is still unsatisfactory. In recent years, many studies have shown that the lncRNA were closely associated with the prognosis of EC patients. In this study, we established a hypoxia-related lncRNA signature, which can effectively distinguish EC patients and predict their survival. To the best of our knowledge, this is the first study to investigate the association of hypoxia-related lncRNA with prognostic features in EC patients.

The hypoxic microenvironment common to cancer cells emerges as an important factor for cancer progression[6, 7]. Besides, hypoxia is an aggressive feature in EC, which can be used for treatments targeting biological changes related to hypoxia[32]. Growing evidence indicates a close correlation between lncRNA and EC[19]. However, there is no systematic study on the correlation between hypoxia and EC. The rapid development of high-throughput gene sequencing technology has laid a foundation for the research of large cancer data[33]. A large number of genomic data were extracted from a single specimen to identify new prognostic and pharmacological biomarkers[34, 35]. This selected risk prediction may be a more targeted and effective prognostic assessment for predicting positive clinical outcomes. Compared with other known prognostic indicators, it may be a more effective classification tool for EC patients.

In the current study, we built a hypoxia-related lncRNA signature. We found that age, grade, stage, disease type and risk score were correlated with the OS of EC patients, and were independent risk factors for the prognosis of EC patients. In order to understand the mechanism of these signature, we used GSEA to analyze the KEGG pathway between different risk groups. We found that the most significant enrichment pathways in the high-risk group were cell cycle and wnt signaling, which were reported associated with the poor prognosis of EC. This partly explained the molecular mechanism of the difference in prognosis between high-risk group and low-risk group. Besides, using CIBERSORT analysis, we found a different immune cell environment characterization of EC between different cluster and risk group. In addition, the risk score was closely related to immunotherapy response, microsatellite instability and TMB.

In order to further confirm the model built in bioinformatics analysis, we select SOS1-IT1 to validate its role in EC cell line. SOS1-IT1 was first reported as the core lncRNA in the prognostic model of ivermectin-related three-lncRNA signature in ovarian cancer[36]. It was also identified as autophagy-related lncRNA

in endometrial cancer[37]. However, its detailed function and molecular mechanism in EC need to be further elucidated. In this study, we found SOS1-IT1 was overexpressed in tumor tissues, and closely correlated with clinicopathological parameters of EC, including the age, disease type, fustat and grade. Moreover, we detected SOS1-IT1 regulated EC cell growth by using knocked down and overexpression experiments. Interestingly, we also found SOS1-IT1 was a functional effector of hypoxia in EC progression. Its expression level was increased in hypoxia or its chemical inducer CoCl_2 . Additionally, the important hypoxia regulatory factor HIF-1 α can directly bind SOS1-IT1 promoter region, and affect its expression level. These results further indicate the relationship between SOS1-IT1 and hypoxia, which may increase EC aggressiveness.

Conclusion

In conclusion, we first clarified the clinical significance of hypoxia-related lncRNA signature in predicting the overall survival rate of EC patients. The hypoxia-related lncRNA was involved in the growth and progression of EC through different pathways. Consistent with previous studies, our findings suggested that some hypoxia-related lncRNA can predict survival outcomes and monitor tumor progression, which increase the reliability of hypoxia-related lncRNA signature. It is also of great significance to reveal the potential molecular mechanism and roles of these lncRNA in other types of malignant tumors.

Abbreviations

EC: Endometrial cancer; OS: Overall survival; ROC: Receiver operating characteristic; GESA: Gene Set Enrichment Analysis, GO: Gene Ontology; KEEG: Kyoto Encyclopedia of Genes and Genomes.

Declarations

Acknowledgements

Not applicable.

Funding

This study was supported by grants from the National Natural Science Foundation of China grants (grant no. 81702780), the Science and Technology Project of Henan province of China (Grant No. 202102310112), the Medical Science and Technology Project of Henan province of China (Grant No. LHGJ20190354).

Author contributions

HYL, JHW, SHC contributed to the manuscript preparation and design of the study. QLF, HYL, JHW carried out the experiments and data analysis. JYL, JL provided their expertise. All authors reviewed the draft and approved the submitted version.

Data availability statement

The data and materials can be found from the first author and corresponding author.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of The Third Affiliated Hospital of Zhengzhou University.

Consent for publication

All listed authors have actively participated in the study and have read and approved the submitted manuscript.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7–30.
2. Raglan O, Kalliala I, Markozannes G, Cividini S, Gunter MJ, Nautiyal J, Gabra H, Paraskevidis E, Martin-Hirsch P, Tsilidis KK, et al. Risk factors for endometrial cancer: An umbrella review of the literature. *Int J Cancer.* 2019;145(7):1719–30.
3. Tsikouras P, Bouchlariotou S, Vrachnis N, Dafopoulos A, Galazios G, Csorba R von. TG: Endometrial cancer: molecular and therapeutic aspects. *Eur J Obstet Gynecol Reprod Biol.* 2013;169(1):1–9.
4. McDonald M, Bender D. Endometrial Cancer: Obesity, Genetics, and Targeted Agents. *Obstet Gynecol Clin N Am.* 2019;46(1):89–105.
5. Habiba M, Pluchino N, Petignat P, Bianchi P, Brosens IA, Benagiano G. Adenomyosis and Endometrial Cancer: Literature Review. *Gynecol Obstet Invest.* 2018;83(4):313–28.
6. Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, Shu Y. Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol Cancer.* 2019;18(1):157.
7. Yang G, Shi R, Zhang Q. Hypoxia and Oxygen-Sensing Signaling in Gene Regulation and Cancer Progression. *Int J Mol Sci.* 2020;21(21):8162.
8. Yuen VW, Wong CC. Hypoxia-inducible factors and innate immunity in liver cancer. *J Clin Invest.* 2020;130(10):5052–62.
9. Hajizadeh F, Okoye I, Esmaily M, Ghasemi Chaleshtari M, Masjedi A, Azizi G, Irandoust M, Ghalamfarsa G, Jadidi-Niaragh F. Hypoxia inducible factors in the tumor microenvironment as therapeutic targets of cancer stem cells. *Life Sci.* 2019;237:116952.

10. Choudhry H, Harris AL. Advances in Hypoxia-Inducible Factor Biology. *Cell Metabol.* 2018;27(2):281–98.
11. Albadari N, Deng S, Li W. The transcriptional factors HIF-1 and HIF-2 and their novel inhibitors in cancer therapy. *Expert opinion on drug discovery.* 2019;14(7):667–82.
12. Semenza GL. HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations. *J Clin Invest.* 2013;123(9):3664–71.
13. Soni S, Padwad YS. HIF-1 in cancer therapy: two decade long story of a transcription factor. *Acta Oncol.* 2017;56(4):503–15.
14. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene.* 2017;36(41):5661–7.
15. Bhan A, Soleimani M, Mandal SS. Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Res.* 2017;77(15):3965–81.
16. Wang J, Zhang X, Chen W, Hu X, Li J, Liu C. Regulatory roles of long noncoding RNAs implicated in cancer hallmarks. *Int J Cancer.* 2020;146(4):906–16.
17. Ding L, Wang R, Shen D, Cheng S, Wang H, Lu Z, Zheng Q, Wang L, Xia L, Li G. Role of noncoding RNA in drug resistance of prostate cancer. *Cell death disease.* 2021;12(6):590.
18. Liu L, Zhang Y, Lu J. The roles of long noncoding RNAs in breast cancer metastasis. *Cell death disease.* 2020;11(9):749.
19. Liu H, Wan J, Chu J. Long non-coding RNAs and endometrial cancer. *Biomed Pharmacother.* 2019;119:109396.
20. Li H, Wang X, Wen C, Huo Z, Wang W, Zhan Q, Cheng D, Chen H, Deng X, Peng C, et al. Long noncoding RNA NORAD, a novel competing endogenous RNA, enhances the hypoxia-induced epithelial-mesenchymal transition to promote metastasis in pancreatic cancer. *Mol Cancer.* 2017;16(1):169.
21. Niu Y, Bao L, Chen Y, Wang C, Luo M, Zhang B, Zhou M, Wang JE, Fang YV, Kumar A, et al. HIF2-Induced Long Noncoding RNA RAB11B-AS1 Promotes Hypoxia-Mediated Angiogenesis and Breast Cancer Metastasis. *Cancer Res.* 2020;80(5):964–75.
22. Shih JW, Chiang WF, Wu ATH, Wu MH, Wang LY, Yu YL, Hung YW, Wang WC, Chu CY, Hung CL, et al. Long noncoding RNA LncHIFCAR/MIR31HG is a HIF-1 α co-activator driving oral cancer progression. *Nat Commun.* 2017;8:15874.
23. Liu H, Zhang Z, Xiong W, Zhang L, Du Y, Liu Y, Xiong X. Long non-coding RNA MALAT1 mediates hypoxia-induced pro-survival autophagy of endometrial stromal cells in endometriosis. *J Cell Mol Med.* 2019;23(1):439–52.
24. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385–95.
25. Wan J, Liu H, Ming L. Lysine crotonylation is involved in hepatocellular carcinoma progression. *Biomed Pharmacother.* 2019;111:976–82.

26. Wan J, Liu H, Yang L, Ma L, Liu J, Ming L. JMJD6 promotes hepatocellular carcinoma carcinogenesis by targeting CDK4. *Int J Cancer*. 2019;144(10):2489–500.
27. Wan J, Liu H, Feng Q, Liu J, Ming L. HOXB9 promotes endometrial cancer progression by targeting E2F3. *Cell death disease*. 2018;9(5):509.
28. Boutros C, Tarhini A, Routier E, Lambotte O, Ladurie FL, Carbonnel F, Izzeddine H, Marabelle A, Champiat S, Berdelou A, et al. Safety profiles of anti-CTLA-4 and anti-PD-1 antibodies alone and in combination. *Nature reviews Clinical oncology*. 2016;13(8):473–86.
29. Suri V, Arora A. Management of Endometrial Cancer: A Review. *Rev Recen Clin Trial*. 2015;10(4):309–16.
30. Moore K, Brewer MA: Endometrial Cancer: Is This a New Disease? *American Society of Clinical Oncology educational book American Society of Clinical Oncology Annual Meeting 2017*, 37:435–442.
31. Morice P, Leary A, Creutzberg C, Abu-Rustum N, Darai E. Endometrial cancer. *Lancet*. 2016;387(10023):1094–108.
32. Salinas-Vera YM, Gallardo-Rincón D, Ruíz-García E, Silva-Cázares MB, de la Peña-Cruz CS, López-Camarillo C. The role of hypoxia in endometrial cancer. *Curr Pharm Biotechnol* 2021(Feb 24):Feb 24.
33. Prokop JW, May T, Strong K, Bilinovich SM, Bupp C, Rajasekaran S, Worthey EA, Lazar J. Genome sequencing in the clinic: the past, present, and future of genomic medicine. *Physiol Genomics*. 2018;50(8):563–79.
34. Ballester LY, Luthra R, Kanagal-Shamanna R, Singh RR. Advances in clinical next-generation sequencing: target enrichment and sequencing technologies. *Expert Rev Mol Diagn*. 2016;16(3):357–72.
35. Su Z, Ning B, Fang H, Hong H, Perkins R, Tong W, Shi L. Next-generation sequencing and its applications in molecular diagnostics. *Expert Rev Mol Diagn*. 2011;11(3):333–43.
36. Li N, Zhan X. Anti-parasite drug ivermectin can suppress ovarian cancer by regulating lncRNA-EIF4A3-mRNA axes. *The EPMA journal*. 2020;11(2):289–309.
37. Wang Z, Zhang J, Liu Y, Zhao R, Zhou X, Wang H: An Integrated Autophagy-Related Long Noncoding RNA Signature as a Prognostic Biomarker for Human Endometrial Cancer: A Bioinformatics-Based Approach. *BioMed research international* 2020, 2020:5717498.

Figures

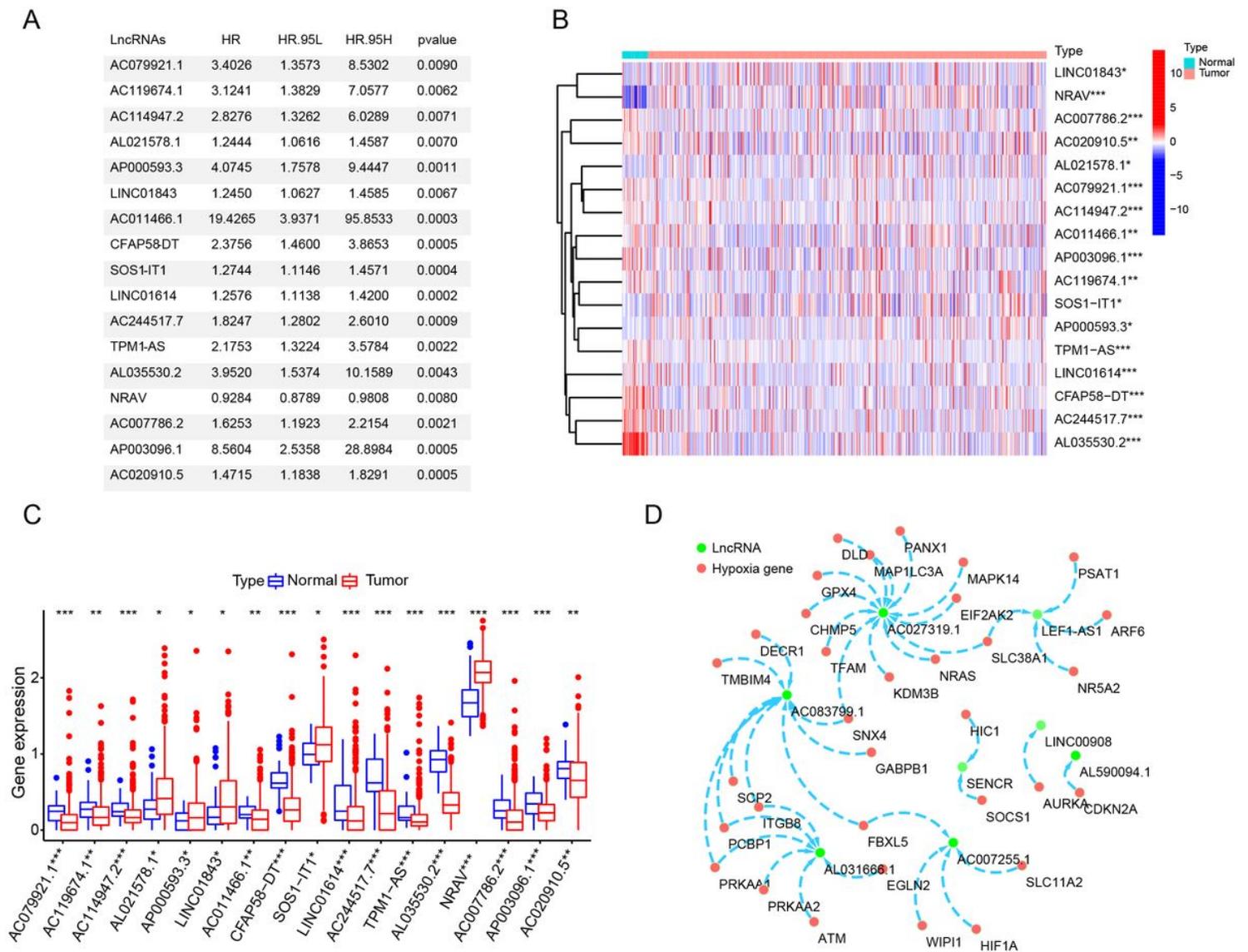


Figure 1

Identification hypoxia-related lncRNA in EC. (A) The univariate Cox regression analysis of 17 hypoxia-related lncRNAs in EC. (B-C) The expression heatmaps (B) and box plot (C) of 17 hypoxia-related lncRNAs in different EC tumor and normal tissue samples. (D) The interaction network based on hypoxia-related lncRNAs and the corresponding hypoxia associated genes. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

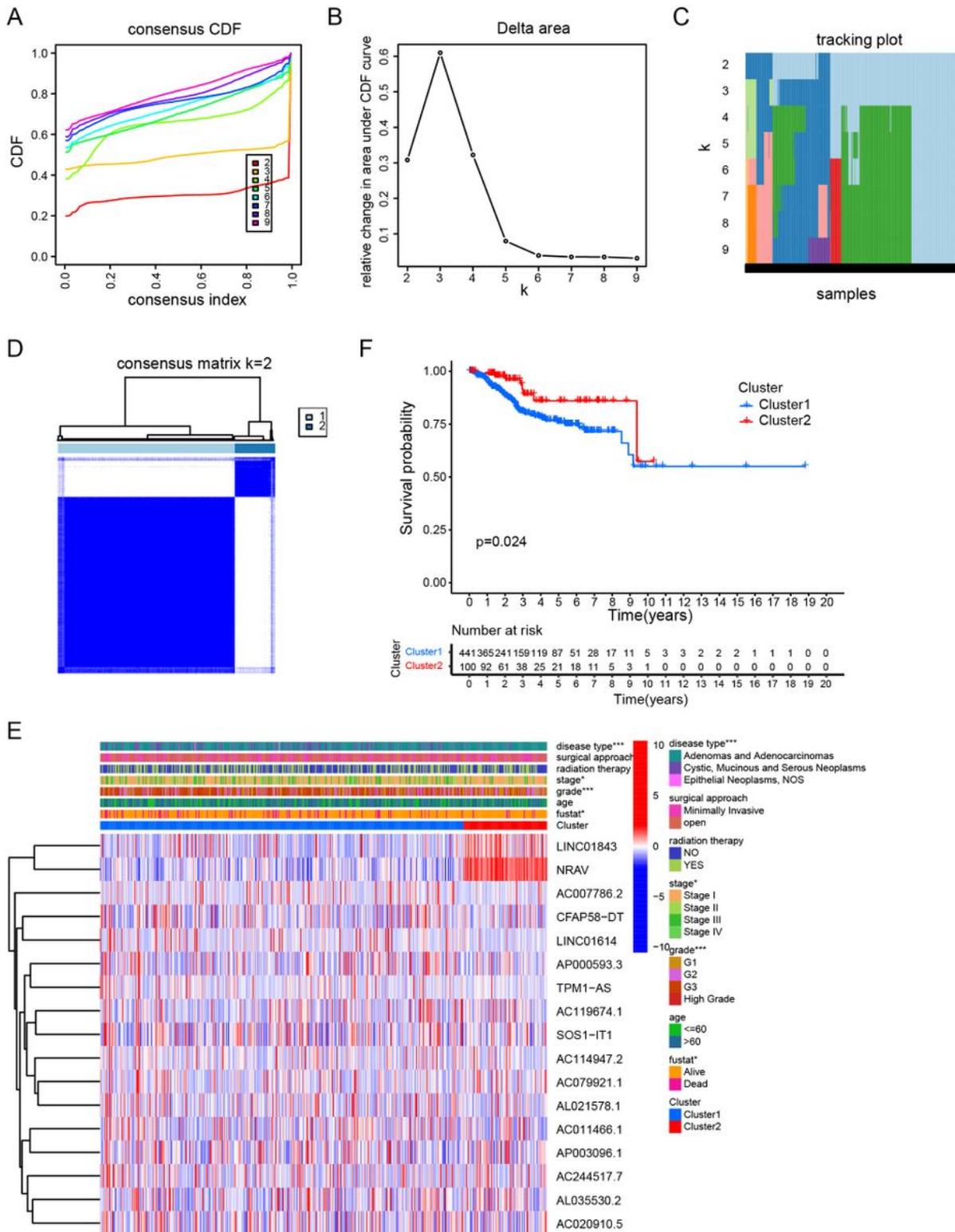


Figure 2

Classification of EC based on hypoxia-related lncRNA. (A) The cumulative distribution function (CDF) plots show the cumulative distributive functions of the consensus matrix for k values (indicated by different colors) between 2 and 9. (B-C) Relative change in area under CDF curve (B) and tracking plot (C) according to various k values. (D) The heatmaps show the consensus clustering matrix for hypoxia-related lncRNA in the EC dataset for k = 2. (E) Heatmap shows the association of clinic pathological

features based on the two clusters. (F) The survival analysis of EC patients between cluster 1 and 2 in TCGA cohort.

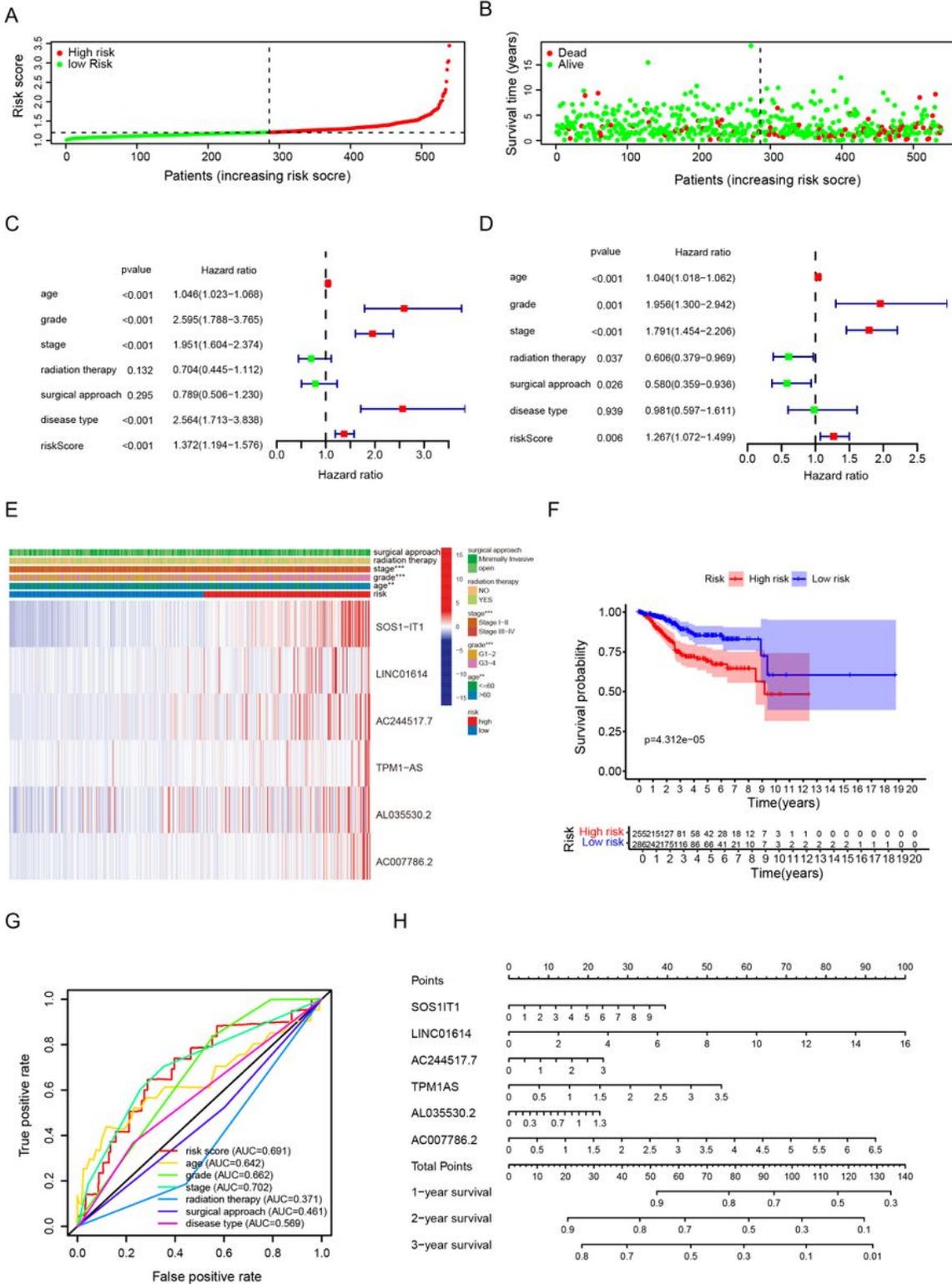


Figure 3

Prognostic value of hypoxia-related lncRNA and construction of a risk signature predicting prognosis. (A) The risk scores for all EC patients in entire cohort are plotted in ascending order, and marked as high-risk (red) and low-risk (blue), as divided by the threshold. (B) Survival days of EC patients were showed in

ascending order of risk parameters. (C-D) The forest plot of univariate (C) and multivariate (D) Cox regression analysis in the cohorts. (E) Heatmap shows the association of risk scores and clinic pathological features based on the six lncRNA signature. (F) The Kaplan-Meier curves of OS in the high-risk and the low-risk groups stratified by the hypoxia-related lncRNA signature in the cohorts. (G) The ROC analysis of OS for the signature and the clinicopathologic parameters. (H) The nomogram for predicting proportion of patients with 1-, 3- or 5-year OS.

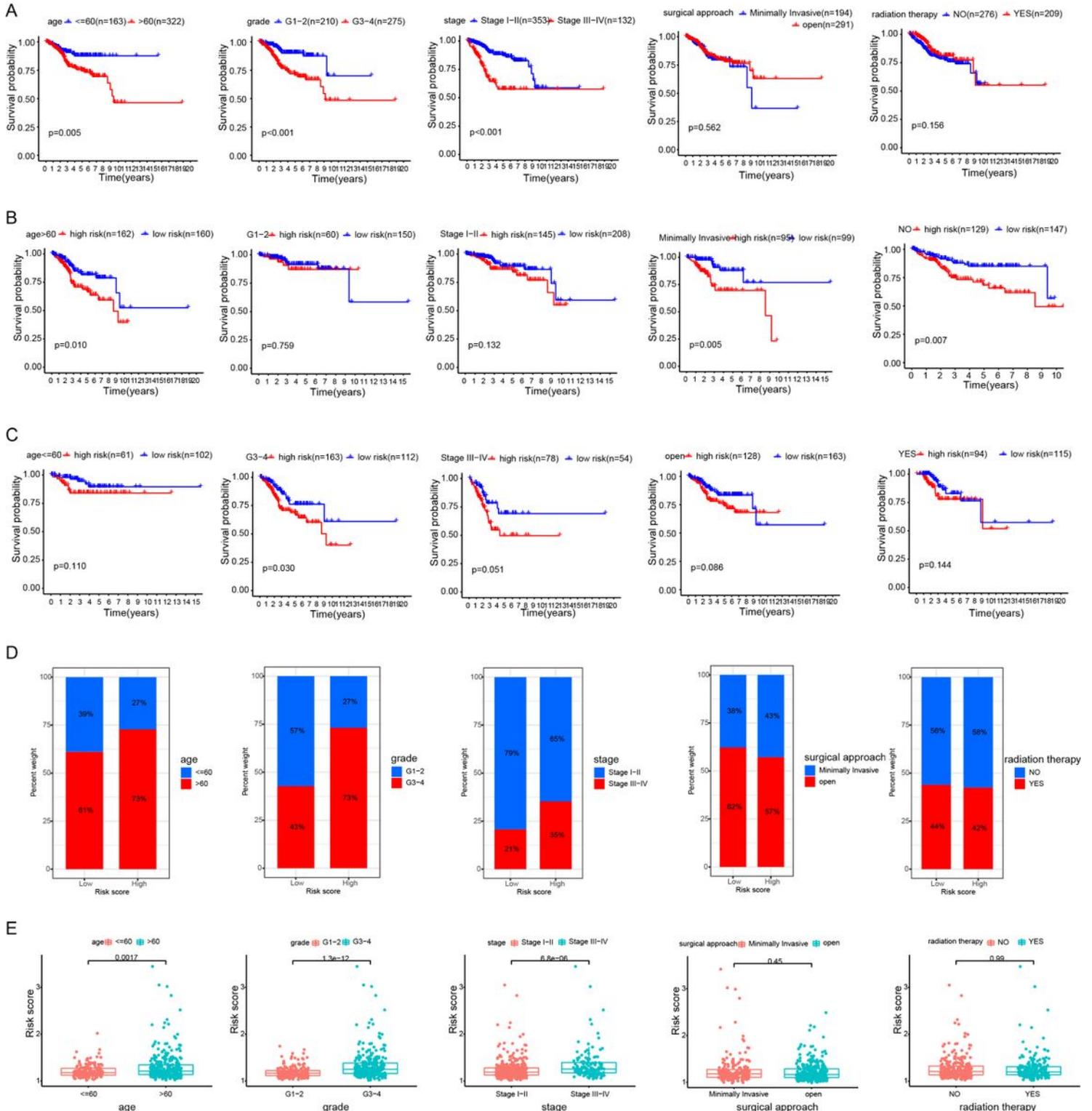


Figure 4

Relationships between the risk score and clinicopathologic factors. (A) The outcome prediction of the hypoxia-related lncRNA signature in patients stratified by age, grade, stage, surgical approach and radiation therapy in TCGA cohort. (B-E) The outcome (B-C) and expression (D-E) in different age (≤ 60 and > 60), grade (G1-G2 and G3-G4), stage (I-II and III-IV), surgical approach (minimally invasive and open) and radiation therapy (No and Yes) group were analyzed.

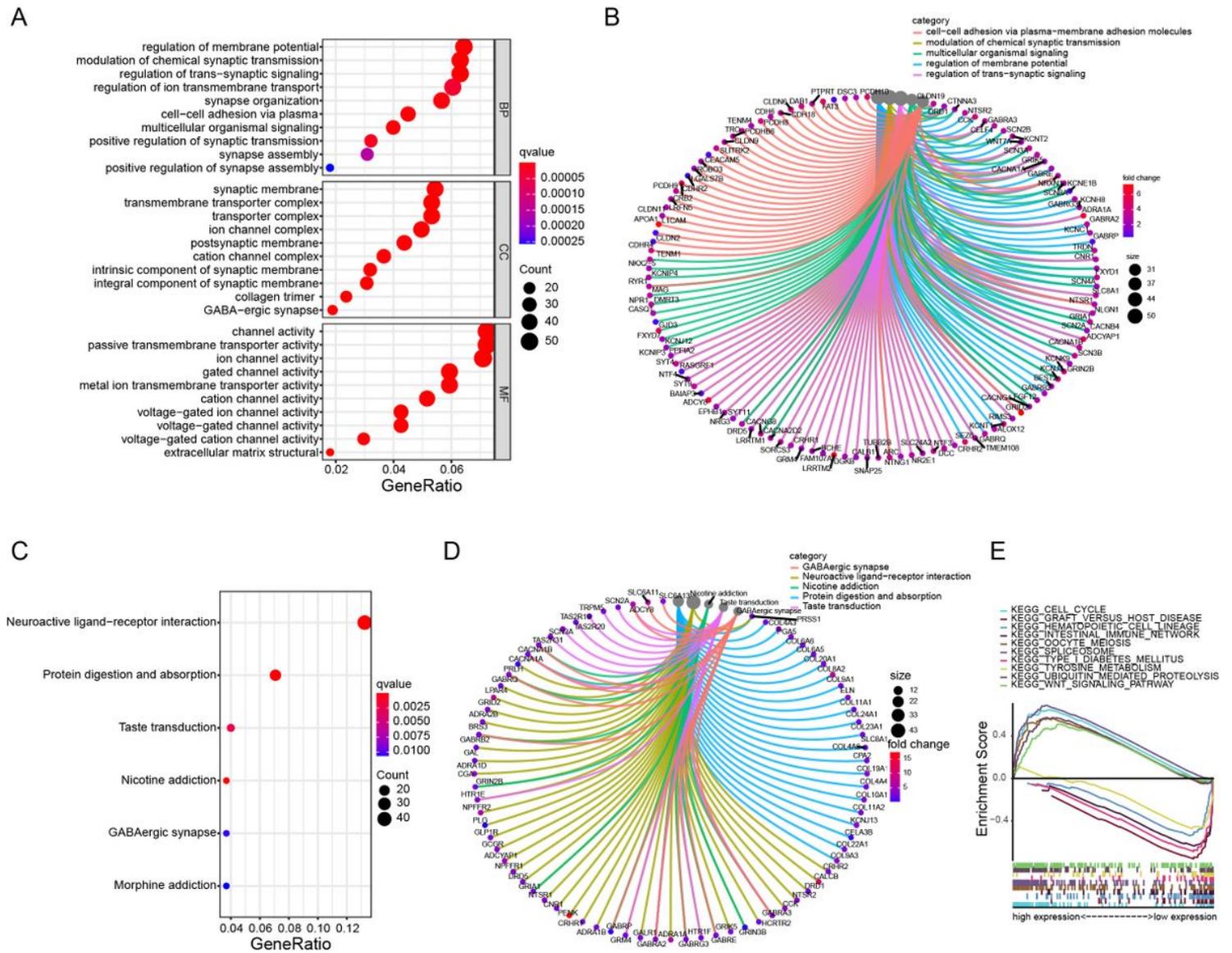


Figure 5

Biological characteristics and pathway analysis of hypoxia-related risky lncRNA. (A-B) Gene Ontology (GO) and (C-D) Kyoto Encyclopedia of Genes and Genomes (KEEG) pathway analysis. (E) The Gene Set Enrichment Analysis (GSEA),

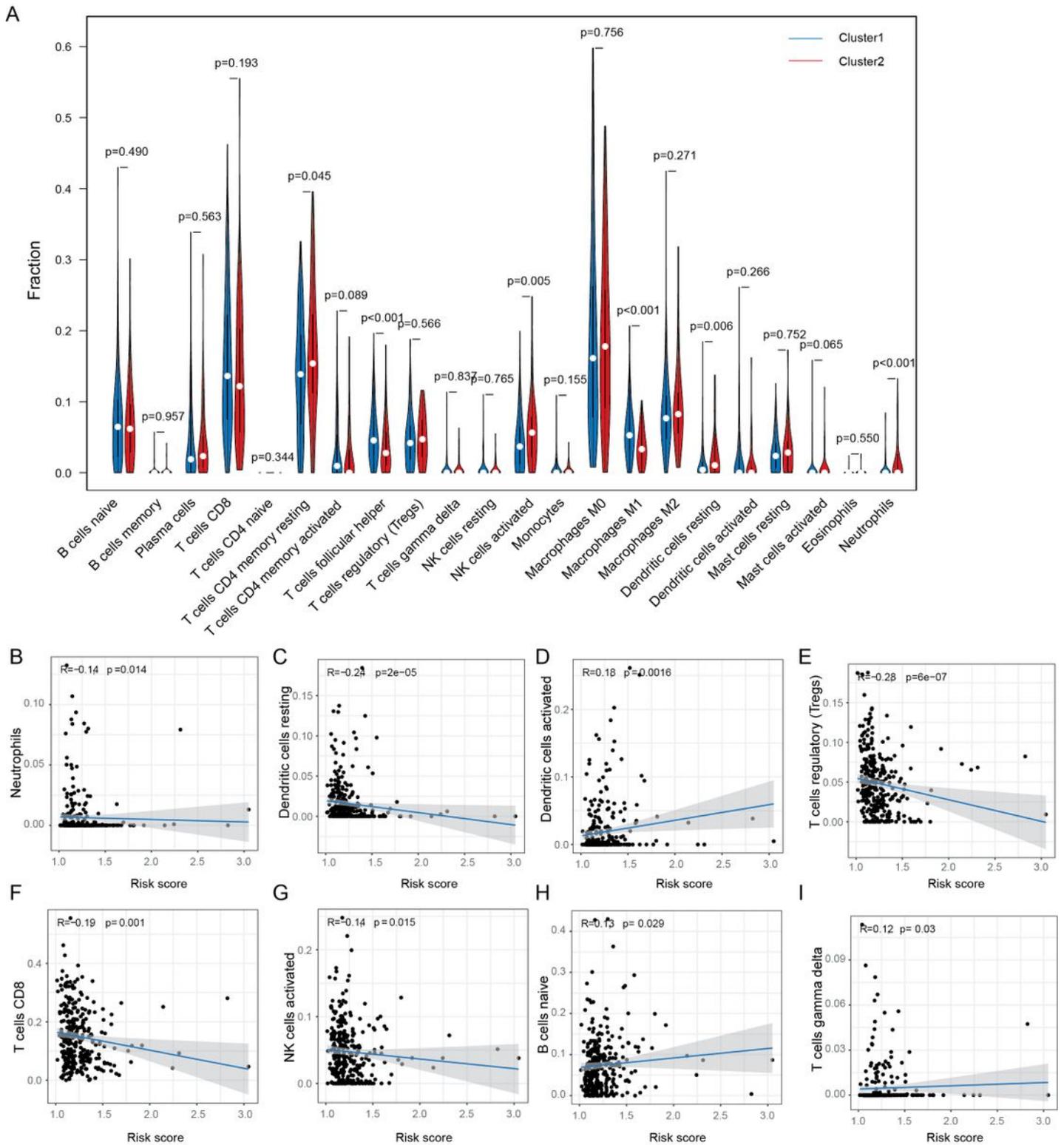


Figure 6

The correlation of risk score with immune cell environment characterization in EC. (A) Violin plots showed the relative levels of immune cells in two clusters. (B-I) The Pearson correlation coefficients between immune cell infiltrating levels and risk scores.

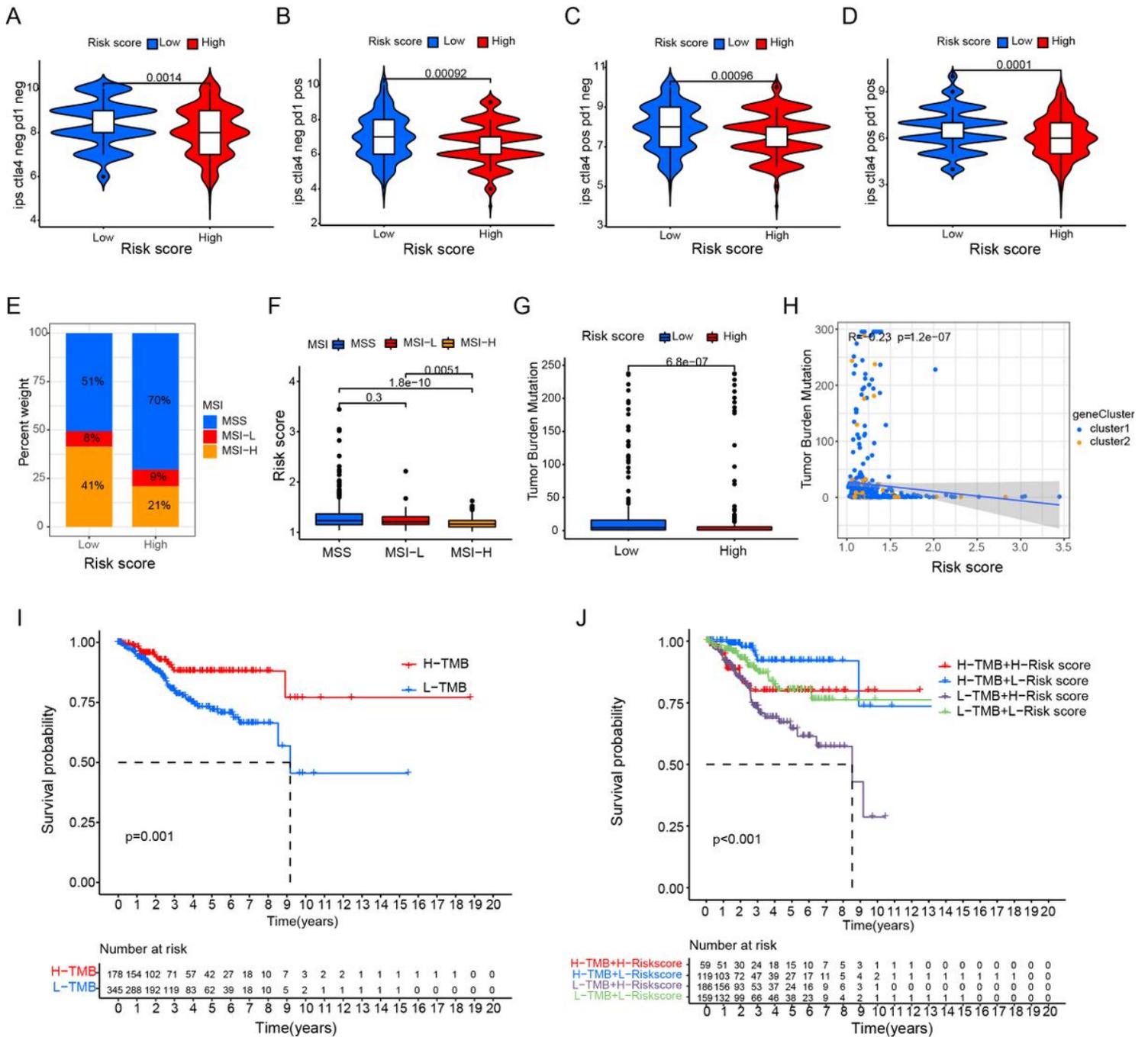


Figure 7

The risk score of hypoxia-related lncRNA was associated with immunotherapy, microsatellite instability and tumor mutation burden in EC. (A-D) The response to anti-CTLA-4 and anti-PD-1 treatment between different risk score group. (E-F) The relationship between hypoxia-related lncRNA risk score and microsatellite instability in EC patients. (G) The tumor mutation burden levels in different risk score group. (H) The Pearson correlation coefficients between tumor mutation burden levels and risk scores. (I-J) Kaplan-Meier analysis showed the correlation between tumor mutation burden and EC prognosis (I), and the difference in survival rate according to risk score between different TMB subgroups (J).

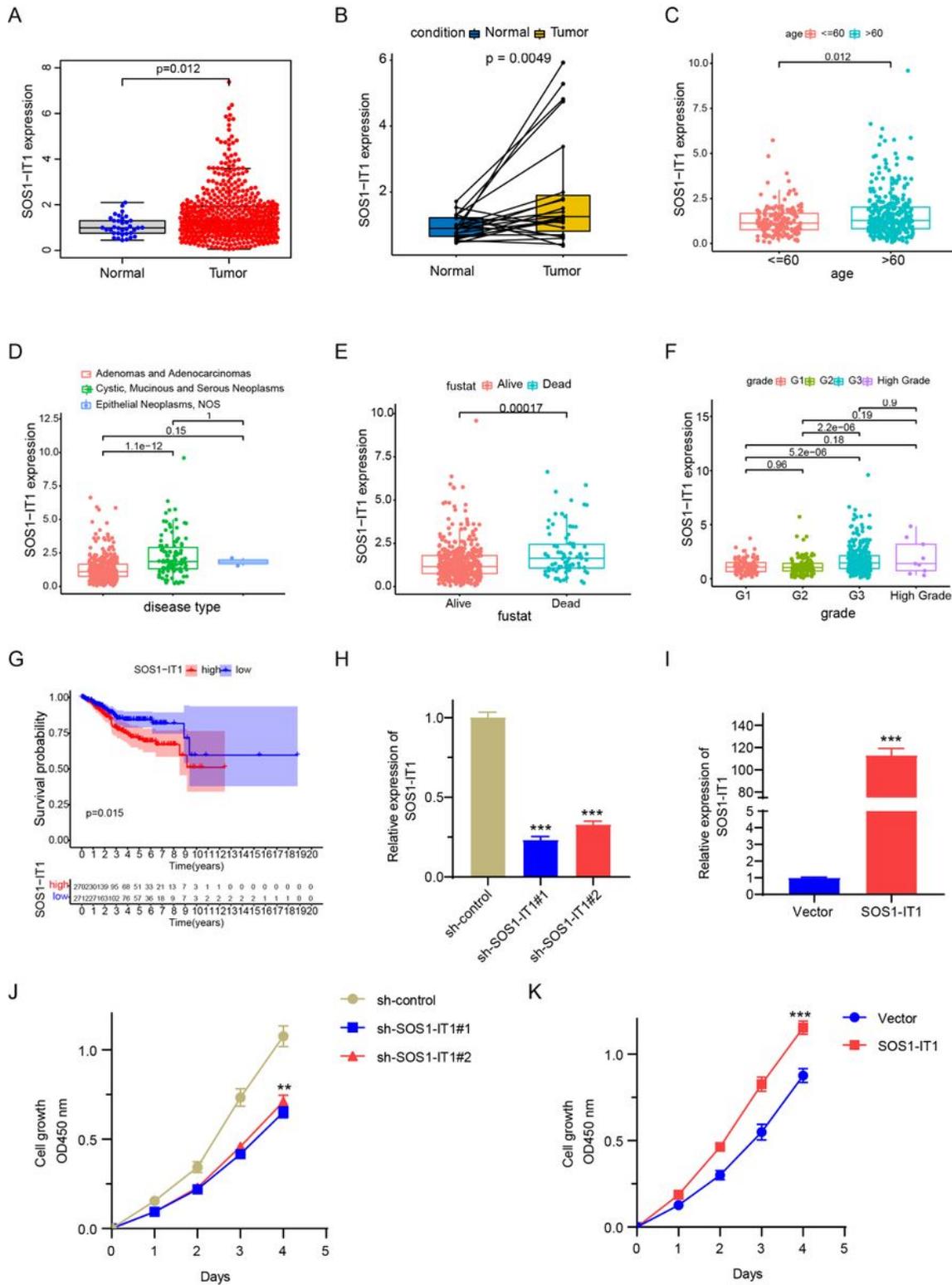


Figure 8

SOS1-IT1 is clinically relevant in EC and promotes EC cell growth. (A-B) The expression level of SOS1-IT1 in normal and tumor tissues of TCGA EC database. (C-F) The correlation analysis of SOS1-IT1 and clinicopathological parameters of EC, including the age, disease type, fustat and grade. (G) Kaplan-Meier analysis of SOS1-IT1 and the prognosis in EC patients. (H-I) The endometrial cancer cell line Ishikawa was knocked down with two different SOS1-IT1 shRNA (H), or transfected with SOS1-IT1 plasmid (I), and

the expression level of SOS1-IT1 was detected by real-time PCR. (J-K) The cell growth was performed by WST-1 assays in SOS1-IT1 knocked down or overexpression in Ishikawa cells. ** P < 0.01, ***P < 0.001.

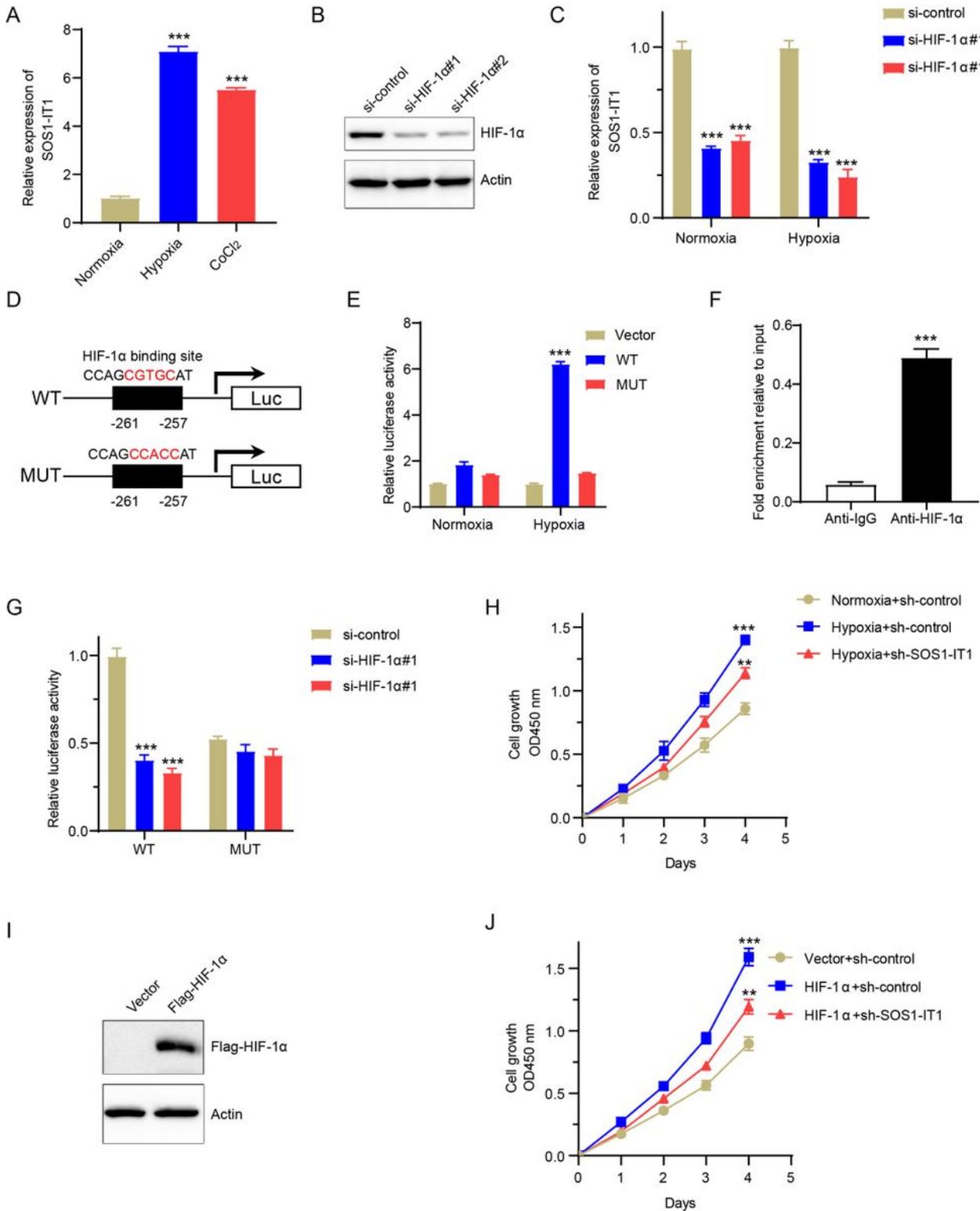


Figure 9

SOS1-IT1 is upregulated under hypoxia and directly transactivated by HIF-1α. (A) The Ishikawa cells were treated with hypoxia or its chemical inducer CoCl₂ for 24 h, and the expression level of SOS1-IT1 was detected by real-time PCR. (B-C) HIF-1α was knocked down by two different HIF-1α small interfering RNA

(siRNA), and the expression level was analyzed by western blot (B). The expression level of SOS1-IT1 was analyzed after knocking down HIF-1 α , both in normoxia and hypoxia condition (C). (D) Schematic illustration of the putative HIF-1 α binding site in SOS1-IT1 gene promoter. (E) The dual-luciferase reporter assays were performed in Ishikawa cells transfected with different reporter constructs, both in normoxia and hypoxia condition. (F) The ChIP assays were performed using control IgG or anti-HIF-1 α antibody in Ishikawa cells. The binding of HIF-1 α on SOS1-IT1 promoter was analyzed by real-time PCR. (G) The dual-luciferase reporter assays were performed in Ishikawa cells when knocking down HIF-1 α with two different siRNAs. (H) The cell growth was performed by WST-1 assays after SOS1-IT1 knocked down in normoxia or hypoxia condition. (I) The Ishikawa cells were overexpressed with flag tagged HIF-1 α , and the expression of HIF-1 α protein was verified by western blot. (J) The cell growth was performed by WST-1 assays after HIF-1 α overexpression with SOS1-IT1 knocked down. ** P < 0.01, ***P < 0.001.