

Integrative Analysis of Genes Associated With Cancer Stem Cell Characteristics via Stemness Indices Combined With Multi-Omics Data in Liver Cancer

Jiecheng Zhang

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

Binhong Yu

Linhai First People's Hospital

Haiwei Chu

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

Bin Hu

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

Luxia Chen

Linhai Second People's Hospital

Weifang Yang

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

Ruili Zhang

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

Wenjun Zhao (✉ 13958561810@163.com)

WenZhou Medical College Affiliated Taizhou Hospital: Taizhou Hospital of Zhejiang Province

<https://orcid.org/0000-0002-8579-3155>

Research

Keywords: Liver cancer, Cancer cell stemness, immune infiltration, ceRNA, prognosis, DNA methylation, copy number variation

Posted Date: November 20th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-111299/v1>

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

[Read Full License](#)

Abstract

Background

Liver cancer (LC) is one of the most prevalent malignancies with a high morbidity and mortality. Cancer stem cells (CSCs) are capable of tumorigenicity, self-renewal and long-term survival. However, the relationship between LC and CSCs remains unclear.

Methods

The clinical data of LC patients and multi-omics data including mRNA-seq, lncRNA-seq, miRNA-seq, mRNA expression-based stemness index (mRNAsi), alternative splicing, DNA methylation and copy number variation (CNV) were obtained from The Cancer Genome Atlas (TCGA). Besides, the clinical data and RNA-seq of validation group were obtained from International Cancer Genome Consortium (ICGC). Integrative analysis of genes associated with CSCs characteristics were performed by stemness indices combined with multi-omics data in LC.

Results

The stemness degree of cells in liver tumor group was higher than that of the normal group. The patients with high stemness degree of tumor tissue had significantly poorer overall survival. Besides, 6 methylation-driven genes and 4 CNV-driven genes related to CSCs characteristics were obtained. Alternative splicing factor regulatory network and ceRNA regulatory network were constructed respectively. Furthermore, the mRNAsi score of LC was negatively correlated with the tumor immune score. The proportion of CD8 T cell, activated memory CD4 T cell, follicular helper T cell and M1 macrophages was up-regulated in tumors with a higher degree of stemness. Finally, a prediction model was established, which has been proved to be a good predictor of prognosis by ICGC data.

Conclusions

Cancer cells with high degree of stemness aggravated the malignant progression of LC, which may be related to low immune infiltration. The proposed prediction model was a powerful predictor for patients and helpful to clinical work.

Background

Liver cancer (LC) is regarded as the sixth most prevalent cancer in the world, ranking fourth among cancer deaths globally, with approximately 840,000 new cases and 780,000 deaths [1]. For males, LC ranks fifth in the incidence of cancer and is the second leading cause of cancer death. For females, it ranks ninth in cancer incidence and sixth in mortality. The major risk factors are characterized by regional distribution, such as Hepatitis B virus (HBV) infection in China, alcoholic consumption and co-infection of HBV and hepatitis C virus (HCV) in Mongolia [2-5]. Surgery is currently an effective treatment for LC but

the postoperative survival rate ranges from 30% to 40% [6]. Therefore, a better understanding of the pathogenesis of LC can provide strategies for the prevention and treatment.

The traditional concept is that all tumor cells within a tumor have tumorigenic ability, but this cannot explain the phenomenon that only a small number of tumor cells can form clones in vitro. Recent studies have found that there is a subpopulation of cells called cancer stem cells (CSCs) within tumors, which has tumorigenic ability and may be the cause of proliferation, recurrence and metastasis [7, 8]. These cells have multiple characteristics, such as self-renewal and drug resistance, promoting tumor escape from treatment [7, 9, 10]. It has been proved that CSCs play a key role in the progress of breast cancer, bladder cancer and lung cancer [11-13]. However, the role of CSCs in LC remains unclear.

The mRNA stemness index (mRNAsi) is a quantitative indicator of the characteristics of CSCs, ranging from 0 to 1. This parameter was established by Malta based on a data set of pluripotent stem cells derived from normal tissue, using machine learning algorithms [14]. Malta has utilized this method to calculate the mRNAsi of all TCGA samples on the basis of the transcriptome data to evaluate the stem cell characteristics [14]. In this study, a comprehensive analysis of the stemness index and multi-omics data from TCGA is conducted to explore the potential impact of CSCs on liver cancer, providing new ideas for the treatment and prognosis evaluation of patients.

Methods

Patient datasets and pre-processing

The data of mRNA-seq (FPKM), lncRNA-seq (FPKM), miRNA-seq, DNA methylation, copy number variation and clinical information were downloaded from TCGA database (<https://portal.gdc.cancer.gov/>). The RNA alternative splicing event data for LC were obtained from the TCGASpliceSeq database (<http://bioinformatics.mdanderson.org/TCGASpliceSeq>). The mRNAsi scores of LC cases in TCGA were available from previous studies [14]. The RNA-seq and clinical information of validation group were downloaded from ICGC. The annotation files of mRNA and lncRNA were obtained from GENCODE (<https://www.genecodegenes.org/>). The annotation file of miRNA was obtained from miRBase (<http://www.mirbase.org/ftp.shtml>).

Correlation analysis between mRNAsi and clinical characteristics

The relationship between mRNAsi and survival was analysed by the survival package in R3.6.0. The relationship between mRNAsi and grades or stages was assessed and visualized by the beeswarm package in R3.6.0.

Screening of differentially expressed genes (DEGs)

After RNA-seq was normalized, the differentially expressed RNAs including mRNA (DEmRNA) and lncRNA (DElncRNA) between the normal group and the tumor group were screened using Wilcox test. The same is true for differentially expressed RNAs between the low and high mRNAsi tumor group. Besides, the edgeR

package was utilized to screen the differentially expressed miRNAs (DEmiRNA). The criterion is $|\log_2 \text{fold change}| > 1$ and adjusted p value < 0.05 . Volcano plots was drawn by the ggplot2 package in R3.6.0.

WGCNA

The WGCNA package in R3.6.0 was used to construct a co-expression network for DEGs [15]. Firstly, the differential gene expression was analyzed, and the hclust function was used to perform sample clustering analysis. Next, the discrete samples were deleted. The function pickSoftThreshold was utilized to establish a weighted adjacency matrix. The soft threshold parameter made the constructed network conform to the power-law distribution, strengthening the strong correlation and weakening the weak correlation. Then we converted the adjacency into a topological overlap matrix (TOM) to measure the network connectivity of genes, and the TOM summed up the adjacent genes for the network gene ratio and calculated the corresponding dissimilarity. We used average linkage hierarchical clustering based on TOM dissimilarity measurement to classify genes showing similar expression profiles with gene modules. The minimum size of the gene group was 10 for the gene dendrogram. We merged some quite similar modules using a cut-off (< 0.25).

Screening of key module genes and correlation analysis

We selected modules most related to the mRNAsi, and genes in these modules were considered to be related to CSCs. GS was defined as correlation between gene expression levels and sample traits; MM was defined as correlation of gene expression between genes in a certain module. For these modules, we calculated GS and MM for each gene. Genes with $MM > 0.8$ and $GS > 0.5$ were selected. The boxplot of key module genes was drawn by ggpubr package in R3.6.0. The Pearson correlations of key module genes were calculated by the corrplot function in R3.6.0.

Functional enrichment analysis for genes and protein–protein interaction (PPI) network analysis

The clusterProfiler package and enrichplot package in R3.6.0 were performed to annotate gene ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) [16]. False discovery rates (FDRs) < 0.05 was considered statistically significant. The PPI network was evaluated in the STRING (<https://string-db.org/>).

Screening of Methylation-driven genes associated with stemness characteristics

We used the normalizeBetweenArrays of limma package in R3.6.0 software for normalization of methylation beta-values. Meanwhile, genes with a value of 0 were excluded. Then, combined with gene expression and methylation data, Methylation-driven genes between the low and high mRNAsi group were selected and drawn by the MethylMix package. The difference in methylation value between the two groups was filtered according to $p < 0.05$. Meanwhile, the correlation coefficient between methylation and gene expression is filtered by less than -0.3 . After the intersection of Methylation-driven genes and DEGs between the two groups, methylation-driven genes associated with stemness characteristics are obtained.

Methylation sites of these genes were extracted. The relationship between Methylation sites and survival was analyzed with the survival package in R3.6.0.

Screening of CNV-driven genes associated with stemness characteristics

The differential CNV genes between the low and high mRNAsi group were selected through Chi square statistic. The location of these genes on the chromosomes was visualized with RCircos package in R3.6.0 software. Combined with gene expression and CNV data, CNV-driven genes between the low and high mRNAsi group were selected through the Kruskal-Wallis test. After the intersection of CNV-driven genes and DEGs between the two groups, CNV-driven genes associated with stemness characteristics are obtained. The relationship between gene expression and CNV was visualized using the beeswarm package in R3.6.0.

Splicing correlation network construction associated with stemness.

Alternative splicing (AS) events of DEGs between the low and high mRNAsi group were visualized using the UpSetR package in R3.6.0. Survival-related AS events was identified by Cox regression analysis with p value less than 0.05. The known splicing factors (SFs) were extracted from the SpliceAid2 database (www.introni.it/spliceaid.html) [17]. The association between survival-related AS events and SFs was further investigated. The regulating event between AS events and SFs was filtered by correlation coefficient more than 0.4. Ultimately, the visualization of network was constructed using Cytoscape (<https://cytoscape.org/>).

Construction of ceRNA network associated with stemness

For DEmRNA and DEmiRNA between the low and high mRNAsi tumor groups, Enrichment analysis of transcription factors was performed with FunRich. TargetScan, miRcode and miRTarBase databases were used to predict the target mRNA of DEmiRNA, and then miRNA-mRNA with intersection in the three databases in were screened. The target miRNA of DElncRNA was predicted on basis of miRcode database. According to the criteria of negative correlation between mRNA and miRNA and positive correlation between mRNA and lncRNA, candidate genes for constructing the ceRNA network were selected. Finally, the visualization of network was constructed using Cytoscape.

Combined analysis of tumor immune microenvironment and mRNAsi

Tumor immune score of each patient was estimated by ESTIMATE package. The correlation coefficient between immune scores and mRNAsi scores were calculated by the corrplot function in R3.6.0. The correlation picture was drawn by ggplot2 package. Besides, CIBERSORT package was utilized to predict the proportion of 22 types of infiltrating immune cells. To improve accuracy, we set Permutations to 1000. Samples were filtered by $p < 0.05$. The significant immune cell type between the low and high mRNAsi tumor group were obtained using Wilcox test.

Construction and validation of immune-related prognostic model

The known immune genes were extracted from the IMMPORT database (<https://www.immport.org/shared/home>). After the intersection of immune genes and DEGs between the low and high mRNAsi tumor groups, differential immune genes were obtained. We deleted patients whose survival time is less than 30. Candidate genes associated with survival were selected by univariate COX analysis and KM survival analysis. Then, prognosis model was established by multivariate analysis. Each patient risk score was calculated based on the expression value of RNA and corresponding coefficients through the formula. Risk score = $\text{Exp1} * \text{Coe1} + \text{Exp2} * \text{Coe2} + \text{Exp3} * \text{Coe3} + \dots + \text{Expn} * \text{Coei}$. The median of the risk score was selected as a cut-of value to divide the patients into two groups, including high-risk group and low-risk group. To evaluate the efficiency of prognostic model, survival receiver operator characteristic curve was applied into the model. Besides, ICGC dataset was used as the validation group and patients with survival time less than 30 were deleted. We calculated the risk values of each patient using the same formula. Then, the patients were divided into two groups using the same cut-of value, including high-risk group and low-risk group. Survival receiver operator characteristic curves were applied into the model.

Univariate and multivariate COX analysis for risk score

The prognostic value of clinical risk factors and risk score were evaluated by univariate and multivariate COX analysis, visualized by the beeswarm package in R3.6.0.

Results

The mRNAsi is higher in LC tissues and correlated with clinical characteristics

In order to explore the potential relationship between mRNAsi and LC, the mRNAsi of tumor tissue was compared with that of normal tissue. The results showed that the mRNAsi in the tumor group was relatively higher than that in the normal group, and also higher than the corresponding paired adjacent tissues (Figure 1A and 1B). When the tumor group was divided into the low mRNAsi group and the high mRNAsi group according to the median value, the results also showed that there was statistical differences between the two groups (Figure 1C). In terms of different grades, the mRNAsi of LC showed significant differences, with a gradually increasing trend (Figure 1D). Besides, the mRNAsi of Stage II and T2 was increased than that of Stage I and T1, respectively (Figure 1E and 1F). In addition, compared with the low mRNAsi group, the patients in the high mRNAsi group had significantly poorer overall survival (Figure 1G).

WGCNA analysis of differential genes between the low and high mRNAsi tumor groups

We noted that mRNAsi was closely related to overall survival, and there was a significant difference in mRNAsi scores between the low and high mRNAsi tumor groups, suggesting that the differential genes between the two groups regulate stem cell characteristics to some extent. Thus, we analyzed the transcriptome data from the two groups. The results revealed that there were 998 DEGs, of which 290 genes were up-regulated and 708 genes were down-regulated (Figure 2A). WGCNA results showed that

the yellow module was highly correlated with mRNAsi (Figure 2B and 2C). According to the filtering criteria of $\text{cor.MM} > 0.8$ and $\text{cor.GS} > 0.5$, 32 key genes of yellow module were screened out (Figure 2D). The heatmap and box plot were pictured based on the expression values of the module genes. The results displayed that the key genes of the yellow module were significantly down-regulated in LC tissues (Figure 2E and Figure 3A).

GO analysis disclosed that these genes were enriched in extracellular structures and angiogenesis (Figure 3B). In terms of KEGG analysis, the results revealed that these key genes mainly focused on PI3K-AKT signaling pathway, ECM- receptor interaction and focal adhesion (Figure 3C). Furthermore, the correlation analysis of these genes showed that there was a strong correlation between the key genes (Figure 3D). As shown in Figure 3E, there is a close interaction between key genes at the protein level. Among them, the edge number of *COL4A4*, *FBN1*, *DCN*, *LUM*, *SPARC* and *THBS2* is more than 7, indicating that these genes are the core genes in the network (Figure 3F).

Screening and analysis of methylation-driven genes

We conducted a combined analysis of DNA methylation and gene expression profiles in the low and high mRNAsi group, and screened out 43 methylation-driven genes (Figure 4A-C). After the intersection of these 43 genes and 998 DEGs, 6 methylation-driven genes are obtained, namely *AQP1*, *CKMT2*, *DCAF4L2*, *GRHL2*, *RAB25*, *SPINT2*. Among them, *DCAF4L2* is hypomethylation, while the rest are hypermethylation. Then, we extracted the values of the methylation sites of these genes, conducted correlation analysis of each methylation site in combination with gene expression profiles, and finally screened the methylation sites according to the correlation coefficient less than -0.3. There were 67 methylation sites that met the conditions, among which *AQP1* accounted for 10, *CKMT2* for 11, *DCAF4L2* for 13, *GRHL2* for 18, *RAB25* for 4, and *SPINT2* for 11 (Figure 4D-I and Supplementary table 4). Additionally, univariate COX regression and KM survival analysis for these 67 methylation sites showed that *CKMT2|cg26206748*, *GRHL2|cg18899777* and *GRHL2|cg18899777* were related to survival (Figure 4J-L).

Screening and analysis of CNV-driven genes

After comparing the CNV values of the low and high mRNAsi groups, 234 differential copy number genes were screened, which were mainly concentrated on chromosome 16 and chromosome 17 (Figure 5A). Subsequently, 181 CNV-driven genes were selected from the 234 genes through the combined analysis of CNV and gene expression profiles. After the intersection of these 181 genes and 998 DEGs, 4 CNV-driven genes are obtained, namely *ABR*, *EFNB3*, *FOXF1*, *ZFP3*, whose number variation types were mostly concentrated in the loss of one copy number and mostly occurred in the high mRNAsi group. As expected, the gene expression of these four genes decreased with the loss of copy number. (Figure 5B-E).

Analysis of survival-associated alternative splicing of DEGs

Among the 998 DEGs, 217 DEGs had alternative splicing events, of which AT events occurred the most, followed by ES events (Figure 6A). Univariate COX regression analysis showed that 58 alternative splicing

events in 217 DEGs were associated with survival (Figure 6B-E). After extracting the gene expression of splicing factors, correlation analysis was performed on the 58 events to construct a splicing network related to survival (Figure 6F).

Construction of lncRNA-miRNA-mRNA ceRNA network related to tumor stemness

803 DEmRNA and 133 DEmiRNA were obtained from the low and high mRNAsi tumor groups. Enrichment analysis of transcription factors using FunRich showed that DEmRNA transcription factors were mainly enriched in *NFIC*, *MYFS*, *FOSB*, *FOS*, *JUND*, *JUNB*, *JUN*, *TEAD1*, *RREB1* and *SRF*, while DEmiRNA transcription factors were chiefly concentrated in *FOXJ2*, *RORA*, *NKX6-1*, *FOXA1*, *NFIC*, *MEF2A*, *SP4*, *POU2F1*, *SP1* and *EGR1* (Figure 7A-B). The results of pathway analysis showed that DEmiRNA was mainly involved in Glypican pathway, proteoglycan syndecan-mediated signaling pathway, IFN-gamma pathway, VEGF and VEGFR signaling network, etc (Figure 7C). GO analysis results presented that DEmiRNA participated in cell communication, signal transduction and GTPase activity, etc (Figure 7D-F). Subsequently, 9 lncRNAs, 5 miRNAs and 70 mRNAs were selected based on TargetScan, miRcode and miRTarBase databases (Figure 7G).

Stemness in tumor cells is related to tumor microenvironment in LC

We evaluated immune score by ESTIMATE method. Interestingly, we observed a phenomenon that the mRNAsi score was slightly negatively correlated with the immune score (Figure 7H). Meanwhile, patients with high mRNAsi scores and low immune scores have a poor prognosis (Figure 7I).

In order to further explore the relationship between immune infiltration and stemness in tumor cells, the proportion of 22 immune cells was evaluated based on CIBERSORT. Figure 8A summarizes the results obtained from 45 tumor samples ($p < 0.05$). Compared with the low mRNAsi tumor group, the high group had higher proportion of CD8⁺ T cell, memory activated CD4⁺ T cell and follicular helper T cell, while naïve B cell, memory resting CD4⁺ T cell and Tregs decreased (Figure 8B). Besides, the mRNAsi score was positively correlated with the proportion of CD8⁺ T cell, memory activated CD4⁺ T cell and follicular helper T cell, while negatively correlated with naïve B cell, memory resting CD4⁺ T cell and Tregs (Figure 8C-G).

Validation of the relationship between stemness and tumor microenvironment.

To systematically understand the relationship between stemness in tumor cells and tumor microenvironment, we evaluated immune score and the proportion of 22 immune cells on 25 types of tumors, including gastric carcinoma, esophageal carcinoma and colon cancer. For immune environment, the immune scores of 13 tumors were significantly negatively correlated with the mRNAsi scores, with brain lower grade glioma being the most correlated (Figure 9A-B). Even though this phenomenon was not evident in the remaining tumors, a negative correlation trend was observed in these tumors (Figure 9A).

Additionally, Figure 9C summarized the results of differential analysis of 22 immune cells from 25 types of tumor ($p < 0.05$). Among the 25 types of tumors, about half of them had apparent differences in the

proportion of CD8+ T cell, resting memory CD4+ T cell, activated memory CD4+ T cell, follicular helper T cell and M1 macrophages. Moreover, the proportion of CD8+ T cell, activated memory CD4+ T cell, follicular helper T cell and M1 macrophages in the high mRNAsi group was mostly up-regulated, while the proportion of resting memory CD4+ T cell was mainly down-regulated. Besides, about one-third of the tumors had obvious contrasts in the proportion of naïve B cell, Tregs and resting Mast cells, which was mainly down-regulated in the high mRNAsi group.

Construction and validation of prognosis model based on immune genes

After the intersection of 1789 immune genes and 998 DEGs, 104 differential immune genes were obtained (Figure 10A). Then, 8 candidate genes were selected following univariate COX analysis and KM survival analysis, which were used for multivariate COX analysis to construct a prognosis model for LC. As shown in figure 10B, each gene in the model was closely associated with survival. Moreover, the prognosis model performed reasonably well in predicting between good and poor prognosis for LC patients (Figure 10C). To further evaluate the efficiency of prognostic model, survival receiver operator characteristic curves were applied into the model, showing that the AUC values of 1-year, 3-year and 5-year were 0.749, 0.750, 0.649, respectively (Figure 10E). The result of risk survival analysis of the ICGC dataset also showed that the prognosis model could clearly distinguish the prognosis of LC patients (Figure 10D). The AUC values of 1-year, 3-year and 5-year were 0.774, 0.802, 0.776, respectively (Figure 10F). To further evaluate the efficiency of prognostic models, we stratified patients using clinical factors. Similarly, the prognosis model still demonstrated a perfect ability to predict the prognosis (Figure 11A-G).

The detailed information between risk value, survival status and model genes is visualized as Figure 12A-D. The prognostic value of clinical risk factors and risk score were evaluated by univariate and multivariate COX analysis. The results showed that pathological T and M stage, pathological stage and riskscore were important prognostic factors in univariate COX analysis, while riskscore was an independent prognostic parameter in multivariate Cox regression. In summary, the prognosis model does performed reasonably well in predicting outcomes in LC patients (Figure 12E-F).

The relationship between prognosis model gene CNV and immune cells

We assessed the relationship between CNV and immune cells on basis of TIMER website (<https://cistrome.shinyapps.io/timer/>). As shown in Figure 12G-K, arm-level gain of *FABP6* and high amplification of *BRIC5* and *FABP6* were related to CD4+ T cell, Macrophage and Neutrophil. Moreover, high amplification of *FABP6* was also associated with B cell, CD8+ T cell and dendritic cell. Besides, arm-level gain of *IL17R* and High amplification of *EPO* and *OXTR* were connected with CD4+ T cell, Macrophage and CD8+ T cell, respectively (Figure 12G-K).

Discussion

LC is regarded as the sixth most prevalent cancer in the world, ranking fourth among cancer deaths globally, with approximately 840,000 new cases and 780,000 deaths [1]. Surgery is currently an effective

treatment for LC, but the postoperative survival rate ranges from 30% to 40% [6]. The concept of CSCs stems from the presence of cells with different characteristics within the tumor [18]. A large body of evidence confirms that there are indeed some cells in the tumor that are capable of self-renewal, tumorigenicity and long-term survival [7, 9, 10]. However, the role of CSCs in LC remains unclear. Therefore, a better understanding of the internal relationship between the two can provide a new strategy for the prevention and treatment of LC.

In this study, the mRNAsi score is regarded as a quantitative indicator of the stemness degree of tumor cells. We first analyzed the difference of the mRNAsi score between normal and tumor tissue, proving that the stemness degree of tumor tissues was higher than that of normal tissues. In terms of different grades, the mRNAsi of HCC had significant differences, showing a gradually increasing trend. Besides, the patients with higher degree of stemness had significantly poorer overall survival. These preliminary results suggest that cancer cells with high degree stemness promote the development of LC.

Based on the above results, WGCNA was performed to further analyze the DEGs and key genes with similar function and related to mRNAsi were obtained. The key genes between low and high mRNAsi tumor samples are related to extracellular structures and angiogenesis. Meanwhile, these genes are enriched in ECM-receptor interaction, regulation of actin cytoskeleton and PI3K-AKT signaling pathway. It is widely accepted that PI3K-AKT signaling pathway is closely associated with the occurrence and development of various tumors. It plays a key role in promoting cell proliferation and inhibiting apoptosis by affecting the activation of various downstream effector proteins. Some experiments have shown that *AKT* can bind a variety of proteins following pathway activation, such as *Bad*, *Caspase9*, *mTOR*, *Par-4* and *P21*, thereby promoting tumor cell survival.[19-22]. Therefore, tumor cells with a high degree of stemness are likely to promote cell growth, metastasis and proliferation by regulating the PI3K-AKT signaling pathway.

Methylation is a major epigenetic modification of DNA, regulating gene function. Abnormal DNA methylation runs through the whole process of tumor tumorigenesis and progression, which can be detected in the early stage of tumor formation and late stage of tumor recurrence and metastasis [23]. Generally, hypermethylation leads to a decrease in gene transcription, while hypomethylation does the opposite. The promoter region of some critical genes such as tumor suppressor gene and DNA repair gene will be hypermethylated, resulting in down-regulated gene expression, which contributes to abnormal cell cycle regulation and DNA damage, ultimately promoting the development of tumors. In this study, we conducted a combined analysis of DNA methylation and gene expression profiles in the low and high mRNAsi group, screening out 6 methylation-driven genes of interest, including *SPINT2*, *DCAF4L2*, *CKMT2*, *AQP1*, *GRHL2*, *RAB25*. Reportedly, the methylation of *SPINT2*, *AQP1*, *GRHL2* and *RAB25* genes was associated with tumor growth and invasion [24-28]. For glioma, the methylation rate of *SPINT2* gene increased with grade [24]. This may be explained by the fact that *SPINT2* encode a membrane-anchored protein that inhibits hepatocyte growth factor (*HGF*), a ligand of the MET receptor. Therefore, abnormal methylation of *SPINT2* gene indirectly activates the MET signaling pathway. For

tumor cells with high degree of stemness in LC, *SPINT2* is likely to aggravate the progression of cancer by this way.

CNV is caused by genomic rearrangement and generally refers to the increase or decrease in the copy number of large fragments of the genome with a length of more than 1kb. Numerous studies have proved that pathogenic CNV can cause mental retardation, growth retardation, autism, tumor and other diseases [29-33]. In tumors, most chromosomes are abnormal. For example, the loss or amplification of fragments of chromosome 17 is related to the malignancy and prognosis of lung cancer, oral carcinoma, ovarian carcinoma, breast cancer and liver cancer [34-37]. Interestingly, 181 CNV-driven genes were mainly concentrated on chromosome 16 and chromosome 17. After the intersection of these 181 genes and 998 DEGs, 4 CNV-driven genes are obtained, namely *ABR*, *EFNB3*, *FOXF1*, *ZFP3*, whose expressions all decreased in the high mRNAsi group, possibly due to the loss of fragments. Reportedly, *FOXF1* inhibits the growth of non-small cell lung cancer by activating tumor suppressor p21 [38]. However, its role and molecular mechanism in high stemness cancer cells remain to be further studied.

The immune microenvironment plays a positive and negative role in the occurrence, development and invasion of cancer. In general, the stronger the immune cell invasion in the tumor microenvironment, the more able to inhibit tumor progression [39-42]. There are three processes in the immune response of cancer: elimination, balance and escape. In the elimination phase, the immune system can eliminate most tumor cells with the help of signaling molecules and cytokines [43-46]. Cancer cells make adaptive changes in the balance phase to evade immune surveillance, such as the epithelial-mesenchymal transition, allowing them to survive in a limited environment [47]. During the escape phase, the immune system can help the tumor founding an immune escape mechanism and promote tumor growth and invasion [48, 49]. The results of our study found that the mRNAsi score was negatively correlated with the tumor immune score, which could be proved by the remaining 24 tumors, especially in LGG, KICH and PRAD. Interestingly, about half of the tumors showed significant differences in the proportion of CD8+ T cell, activated memory CD4+ T cell, follicular helper T cell and M1 macrophages, which were up-regulated in the high mRNAsi group. Considering that patients with high mRNAsi scores and low immune scores have a worse prognosis, we hold the opinion that although the proportion of above immune cells is up-regulated in high mRNAsi tumors, the total number of immune cells infiltrated in the tumor is small and the killing effect on the tumor is poor. This may facilitate the process of immune escape for tumor cells with a higher degree of stemness.

Considering that immune infiltration and the stemness degree of tumor cells influence the prognosis of patients, we constructed a prognostic model by combining immune-related genes and stemness-related gene. The efficiency of the model was verified by ICGC dataset and TCGA dataset stratified by clinical factors, showing a perfect ability to predict the prognosis.

Taken together, our current findings demonstrate that the stemness degree of cells in liver tumor group is higher than that of the normal group. Besides, the patients with high stemness degree of cancer cells had significantly poorer overall survival. A combined analysis of multi-omics data for the low and high

mRNAsi tumor samples reveals 6 methylation-driven genes and 4 CNV-driven genes, some of which have been proved to be related to the occurrence and development of tumor. In addition, we construct an alternative splicing factor regulatory network and a ceRNA regulatory network to elucidate the inter-regulatory relationship between genes associated with stemness characteristics of liver cancer cells. We also find that the mRNAsi score of LC is negatively correlated with the tumor immune score, and the proportion of CD8+ T cell, activated memory CD4+ T cell, follicular helper T cell and M1 macrophages are up-regulated in the high mRNAsi group, which are verified in other 24 tumors. Finally, we constructed a prognostic model, which has been proved to be a good predictor of prognosis. The above research results can help us to better understand the mechanism of cancer progression and provide new strategies for the prevention and treatment of liver cancer. Of course, there are some limitations in our study, and the above conclusions need to be verified by further biological studies.

Conclusions

Cancer cells with high degree of stemness aggravated the malignant progression of LC, which may be related to low immune infiltration. The proposed prediction model was a powerful predictor for patients and helpful to clinical work.

Abbreviations

LC: Liver cancer

CSCs: Cancer stem cells

mRNAsi: mRNA expression-based stemness index

CNV: copy number variation

TCGA: The Cancer Genome Atlas

ICGC: International Cancer Genome Consortium

HBV: Hepatitis B virus

HCV: Hepatitis C virus

DEmRNA: The differentially expressed mRNAs

DEmiRNA: The differentially expressed miRNAs

DElncRNA: The differentially expressed lncRNAs

PPI: Protein–protein interaction

AS: Alternative splicing

SFs: splicing factors

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analysed during the current study are available in the TCGA repository:

<https://portal.gdc.cancer.gov/>.

Competing interests

The authors declare that they have no competing interests

Funding

This study was supported by the Natural Science Foundation of Zhejiang (No.LGF20H020007).

Authors' contributions

WZ, RZ, WY: conception of the work. JZ: data analysis and drafting the article. BY, HC, BH, and LC: data collection, data analysis, review and a critical revision of the manuscript.

Acknowledgements

None

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**:394-424.
2. Chimed T, Sandagdorj T, Znaor A, Laversanne M, Tseveen B, Genden P, Bray F: **Cancer incidence and cancer control in Mongolia: Results from the National Cancer Registry 2008-12.** *Int J Cancer* 2017, **140**:302-309.

3. Cui Y, Jia J: **Update on epidemiology of hepatitis B and C in China.***J Gastroenterol Hepatol* 2013, **28 Suppl 1**:7-10.
4. Bruix J, Gores GJ, Mazzaferro V: **Hepatocellular carcinoma: clinical frontiers and perspectives.***Gut* 2014, **63**:844-855.
5. Fung J, Cheung KS, Wong DK, Mak LY, To WP, Seto WK, Lai CL, Yuen MF: **Long-term outcomes and predictive scores for hepatocellular carcinoma and hepatitis B surface antigen seroclearance after hepatitis B e-antigen seroclearance.***Hepatology* 2018, **68**:462-472.
6. Xia L, Huang W, Tian D, Zhu H, Qi X, Chen Z, Zhang Y, Hu H, Fan D, Nie Y, Wu K: **Overexpression of forkhead box C1 promotes tumor metastasis and indicates poor prognosis in hepatocellular carcinoma.***Hepatology* 2013, **57**:610-624.
7. Tang DG: **Understanding cancer stem cell heterogeneity and plasticity.***Cell Res* 2012, **22**:457-472.
8. De Angelis ML, Francescangeli F, La Torre F, Zeuner A: **Stem Cell Plasticity and Dormancy in the Development of Cancer Therapy Resistance.***Front Oncol* 2019, **9**:626.
9. Li L, Neaves WB: **Normal stem cells and cancer stem cells: the niche matters.***Cancer Res* 2006, **66**:4553-4557.
10. Dalerba P, Dylla SJ, Park IK, Liu R, Wang X, Cho RW, Hoey T, Gurney A, Huang EH, Simeone DM, et al: **Phenotypic characterization of human colorectal cancer stem cells.***Proc Natl Acad Sci U S A* 2007, **104**:10158-10163.
11. Pei J, Wang Y, Li Y: **Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis.***J Transl Med* 2020, **18**:74.
12. Pan S, Zhan Y, Chen X, Wu B, Liu B: **Identification of Biomarkers for Controlling Cancer Stem Cell Characteristics in Bladder Cancer by Network Analysis of Transcriptome Data Stemness Indices.***Front Oncol* 2019, **9**:613.
13. Liao Y, Wang Y, Cheng M, Huang C, Fan X: **Weighted Gene Coexpression Network Analysis of Features That Control Cancer Stem Cells Reveals Prognostic Biomarkers in Lung Adenocarcinoma.***Front Genet* 2020, **11**:311.
14. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kaminska B, Huelsken J, Omberg L, Gevaert O, et al: **Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation.***Cell* 2018, **173**:338-354 e315.
15. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.***BMC Bioinformatics* 2008, **9**:559.
16. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.***OMICS* 2012, **16**:284-287.
17. Piva F, Giulietti M, Burini AB, Principato G: **SpliceAid 2: a database of human splicing factors expression data and RNA target motifs.***Hum Mutat* 2012, **33**:81-85.
18. Peitzsch C, Tyutyunnykova A, Pantel K, Dubrovskaya A: **Cancer stem cells: The root of tumor recurrence and metastases.***Semin Cancer Biol* 2017, **44**:10-24.

19. Wanigasooriya K, Tyler R, Barros-Silva JD, Sinha Y, Ismail T, Beggs AD: **Radiosensitising Cancer Using Phosphatidylinositol-3-Kinase (PI3K), Protein Kinase B (AKT) or Mammalian Target of Rapamycin (mTOR) Inhibitors.***Cancers (Basel)* 2020, **12**.
20. Xu F, Na L, Li Y, Chen L: **Roles of the PI3K/AKT/mTOR signalling pathways in neurodegenerative diseases and tumours.***Cell Biosci* 2020, **10**:54.
21. Iida M, Harari PM, Wheeler DL, Toulany M: **Targeting AKT/PKB to improve treatment outcomes for solid tumors.***Mutat Res* 2020, **819-820**:111690.
22. Ortega MA, Fraile-Martinez O, Asunsolo A, Bujan J, Garcia-Honduvilla N, Coca S: **Signal Transduction Pathways in Breast Cancer: The Important Role of PI3K/Akt/mTOR.***J Oncol* 2020, **2020**:9258396.
23. Werner RJ, Kelly AD, Issa JJ: **Epigenetics and Precision Oncology.***Cancer J* 2017, **23**:262-269.
24. Fukushima T, Kawaguchi M, Yamamoto K, Yamashita F, Izumi A, Kaieda T, Takezaki Y, Itoh H, Takeshima H, Kataoka H: **Aberrant methylation and silencing of the SPINT2 gene in high-grade gliomas.***Cancer Sci* 2018, **109**:2970-2979.
25. Wang N, Che Y, Yin F, Yu F, Bi X, Wang Y: **Study on the methylation status of SPINT2 gene and its expression in cervical carcinoma.***Cancer Biomark* 2018, **22**:435-442.
26. Yin X, Huang H, Huang S, Xu A, Fan F, Luo S, Yan H, Chen L, Sun C, Hu Y: **A Novel Scoring System for Risk Assessment of Elderly Patients With Cytogenetically Normal Acute Myeloid Leukemia Based on Expression of Three AQP1 DNA Methylation-Associated Genes.***Front Oncol* 2020, **10**:566.
27. Pistore C, Giannoni E, Colangelo T, Rizzo F, Magnani E, Muccillo L, Giurato G, Mancini M, Rizzo S, Riccardi M, et al: **DNA methylation variations are required for epithelial-to-mesenchymal transition induced by cancer-associated fibroblasts in prostate cancer cells.***Oncogene* 2017, **36**:5551-5566.
28. Gu Y, Zou YM, Lei D, Huang Y, Li W, Mo Z, Hu Y: **Promoter DNA methylation analysis reveals a novel diagnostic CpG-based biomarker and RAB25 hypermethylation in clear cell renal cell carcinoma.***Sci Rep* 2017, **7**:14200.
29. Diskin SJ, Hou C, Glessner JT, Attiyeh EF, Laudenslager M, Bosse K, Cole K, Mosse YP, Wood A, Lynch JE, et al: **Copy number variation at 1q21.1 associated with neuroblastoma.***Nature* 2009, **459**:987-991.
30. Wang J, Wang Y, Wang L, Chen WY, Sheng M: **The diagnostic yield of intellectual disability: combined whole genome low-coverage sequencing and medical exome sequencing.***BMC Med Genomics* 2020, **13**:70.
31. Zhong Q, Fan J, Chu H, Pang M, Li J, Fan Y, Liu P, Wu C, Qiao J, Li R, Hang J: **Integrative analysis of genomic and epigenetic regulation of endometrial cancer.***Aging (Albany NY)* 2020, **12**.
32. Akpa CA, Kleo K, Oker E, Tomaszewski N, Messerschmidt C, Lopez C, Wagener R, Oehl-Huber K, Dettmer K, Schoeler A, et al: **Acquired resistance to DZNep-mediated apoptosis is associated with copy number gains of AHCY in a B-cell lymphoma model.***BMC Cancer* 2020, **20**:427.
33. Ashitha SNM, Ramachandra NB: **Integrated Functional Analysis Implicates Syndromic and Rare Copy Number Variation Genes as Prominent Molecular Players in Pathogenesis of Autism Spectrum Disorders.***Neuroscience* 2020, **438**:25-40.

34. Birkness JE, Spada NG, Miller C, Luketich JD, Nason KS, Sun W, Davison JM: **Extreme chromosome 17 copy number instability is a prognostic factor in patients with gastroesophageal adenocarcinoma: A retrospective cohort study.** 2018, *57*:28-34.
35. Du C, Mark D, Wappenschmidt B, Böckmann B, Pabst B, Chan S, Cao H, Morlot S, Scholz C, Auber B, et al: **A tandem duplication of BRCA1 exons 1-19 through DHX8 exon 2 in four families with hereditary breast and ovarian cancer syndrome.***Breast Cancer Res Treat* 2018, *172*:561-569.
36. Lee K, Jang MH, Chung YR, Lee Y, Kang E, Kim SW, Kim YJ, Kim JH, Kim IA, Park SY: **Prognostic significance of centromere 17 copy number gain in breast cancer depends on breast cancer subtype.***Hum Pathol* 2017, *61*:111-120.
37. Arnold A, Bahra M, Lenze D, Bradtmöller M, Guse K, Gehlhaar C, Bläker H, Heppner FL, Koch A: **Genome wide DNA copy number analysis in cholangiocarcinoma using high resolution molecular inversion probe single nucleotide polymorphism assay.***Exp Mol Pathol* 2015, *99*:344-353.
38. Wu CY, Chan CH, Dubey NK, Wei HJ, Lu JH, Chang CC, Cheng HC, Ou KL, Deng WP: **Highly Expressed FOXF1 Inhibit Non-Small-Cell Lung Cancer Growth via Inducing Tumor Suppressor and G1-Phase Cell-Cycle Arrest.***Int J Mol Sci* 2020, *21*.
39. Best SA, Hess J, Souza-Fonseca-Guimaraes F, Cursons J, Kersbergen A, Dong X, Rautela J, Hyslop SR, Ritchie ME, Davis MJ, et al: **Harnessing natural killer immunity in metastatic small cell lung cancer.***J Thorac Oncol* 2020.
40. Workel HH, van Rooij N, Plat A, Spierings DCJ: **Transcriptional Activity and Stability of CD39+CD103+CD8+ T Cells in Human High-Grade Endometrial Cancer.** 2020, *21*.
41. Sheng H, Huang Y, Xiao Y, Zhu Z, Shen M, Zhou P, Guo Z, Wang J, Wang H, Dai W, et al: **ATR inhibitor AZD6738 enhances the antitumor activity of radiotherapy and immune checkpoint inhibitors by potentiating the tumor immune microenvironment in hepatocellular carcinoma.** 2020, *8*.
42. Tay RE, Richardson EK, Toh HC: **Revisiting the role of CD4(+) T cells in cancer immunotherapy-new insights into old paradigms.***Cancer Gene Ther* 2020.
43. Egeblad M, Nakasone ES, Werb Z: **Tumors as organs: complex tissues that interface with the entire organism.***Dev Cell* 2010, *18*:884-901.
44. Mantovani A: **Molecular pathways linking inflammation and cancer.***Curr Mol Med* 2010, *10*:369-373.
45. Qian BZ, Pollard JW: **Macrophage diversity enhances tumor progression and metastasis.***Cell* 2010, *141*:39-51.
46. Murdoch C, Muthana M, Coffelt SB, Lewis CE: **The role of myeloid cells in the promotion of tumour angiogenesis.***Nat Rev Cancer* 2008, *8*:618-631.
47. Ricciardi M, Zanotto M, Malpeli G, Bassi G, Perbellini O, Chilosi M, Bifari F, Krampera M: **Epithelial-to-mesenchymal transition (EMT) induced by inflammatory priming elicits mesenchymal stromal cell-like immune-modulatory properties in cancer cells.***Br J Cancer* 2015, *112*:1067-1075.
48. Tao M, Liu L, Shen M, Zhi Q, Gong FR, Zhou BP, Wu Y, Liu H, Chen K, Shen B, et al: **Inflammatory stimuli promote growth and invasion of pancreatic cancer cells through NF-κB pathway dependent repression of PP2Ac.***Cell Cycle* 2016, *15*:381-393.

49. Pathak BR, Breed AA, Apte S, Acharya K, Mahale SD: **Cysteine-rich secretory protein 3 plays a role in prostate cancer cell invasion and affects expression of PSA and ANXA1.** *Mol Cell Biochem* 2016, 411:11-21.

Figures

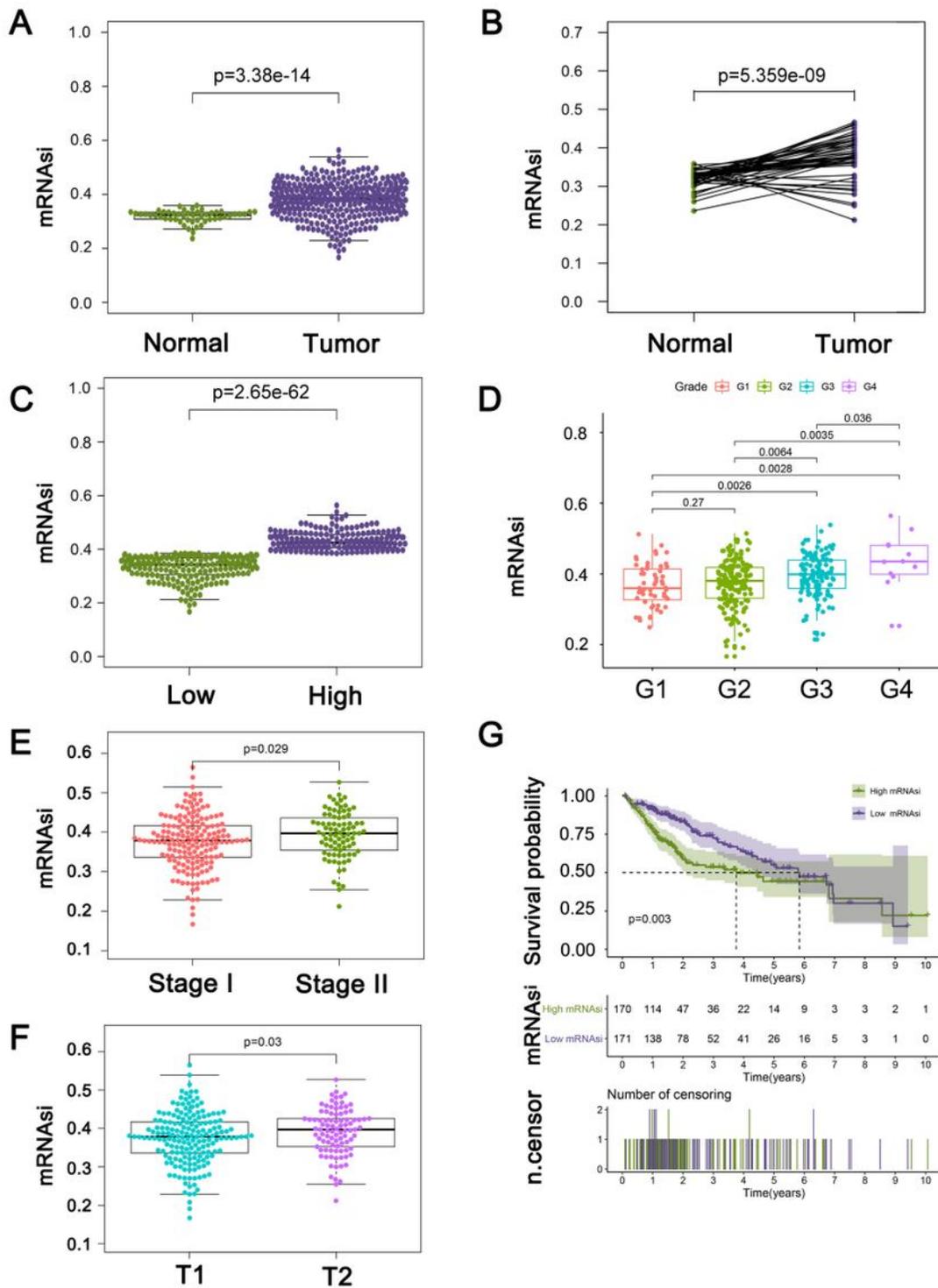


Figure 1

Correlation between mRNAsi and clinical characteristics in LC (A). Differences in mRNAsi between normal (50 samples) and tumor (371 samples) tissues. (B). Differences in mRNAsi between adjacent normal tissue (50 samples) and corresponding tumor (50 samples) tissues. (C). Differences in mRNAsi between the low (186 samples) and high (185 samples) mRNAsi tumor group. (D). Comparison of mRNAsi in different clinical grades. (E). Comparison of mRNAsi between clinical stage I and stage II. (F). Comparison of mRNAsi in different clinical tumour depths. (G). Kaplan–Meier survival curves of mRNAsi in LC.

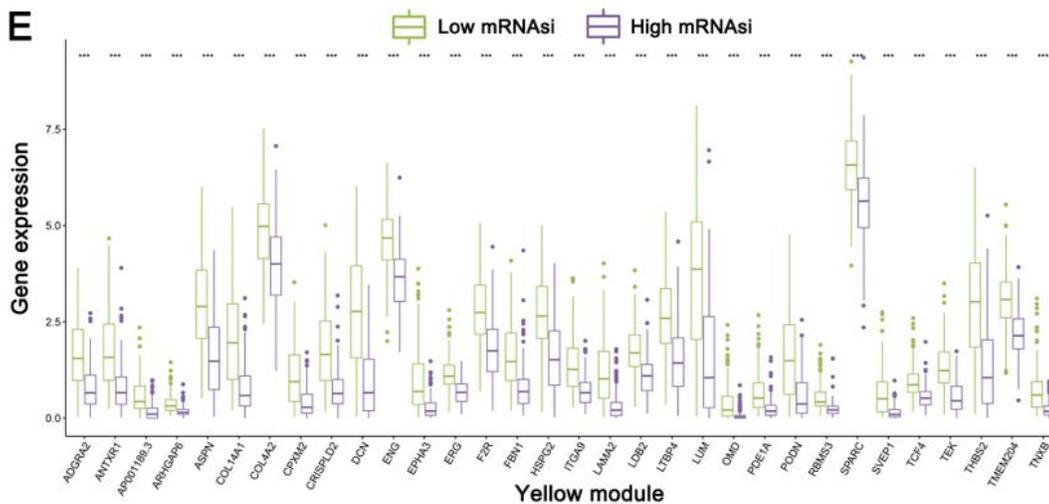
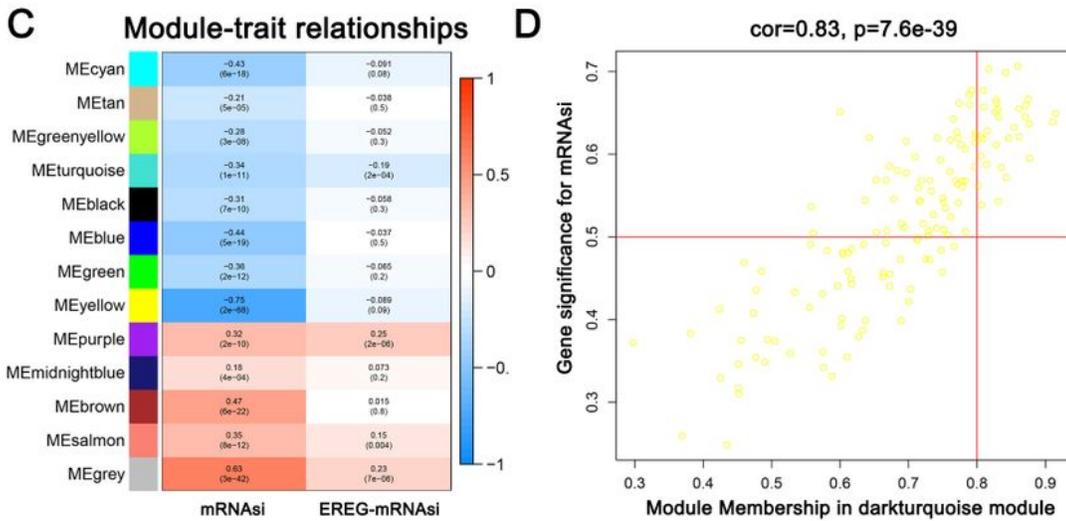
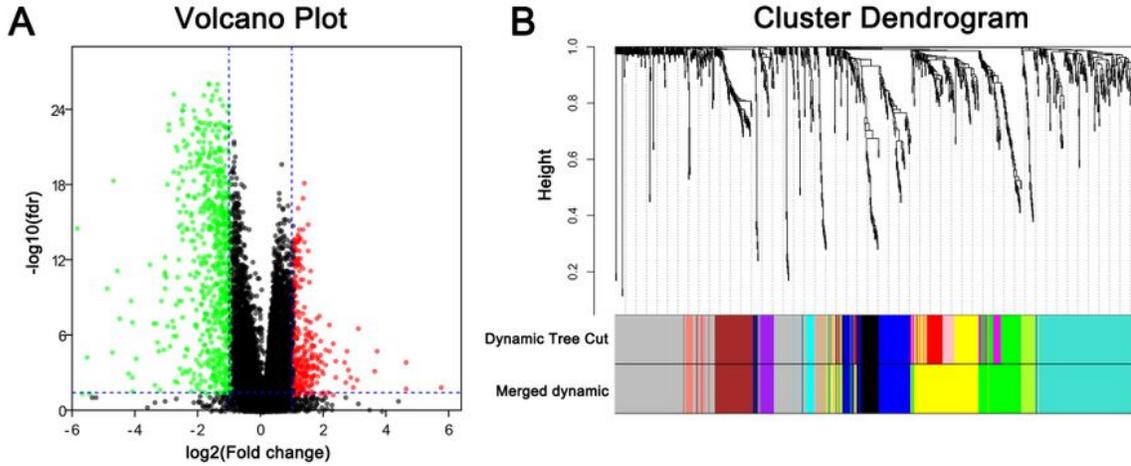


Figure 2

WGCNA analysis between the low and high mRNAsi tumor tissues. (A). Volcano plots of DGEs between low and high mRNAsi tissues. (B). Cluster dendrogram of WGCNA analysis. (C). Module-trait relationships of WGCNA analysis. (D). Module membership in yellow module and gene significance for mRNAsi. (E). The expression value of key module genes. Green indicates low mRNAsi tissues (186 samples) and purple indicates high mRNAsi tissues (185 samples).

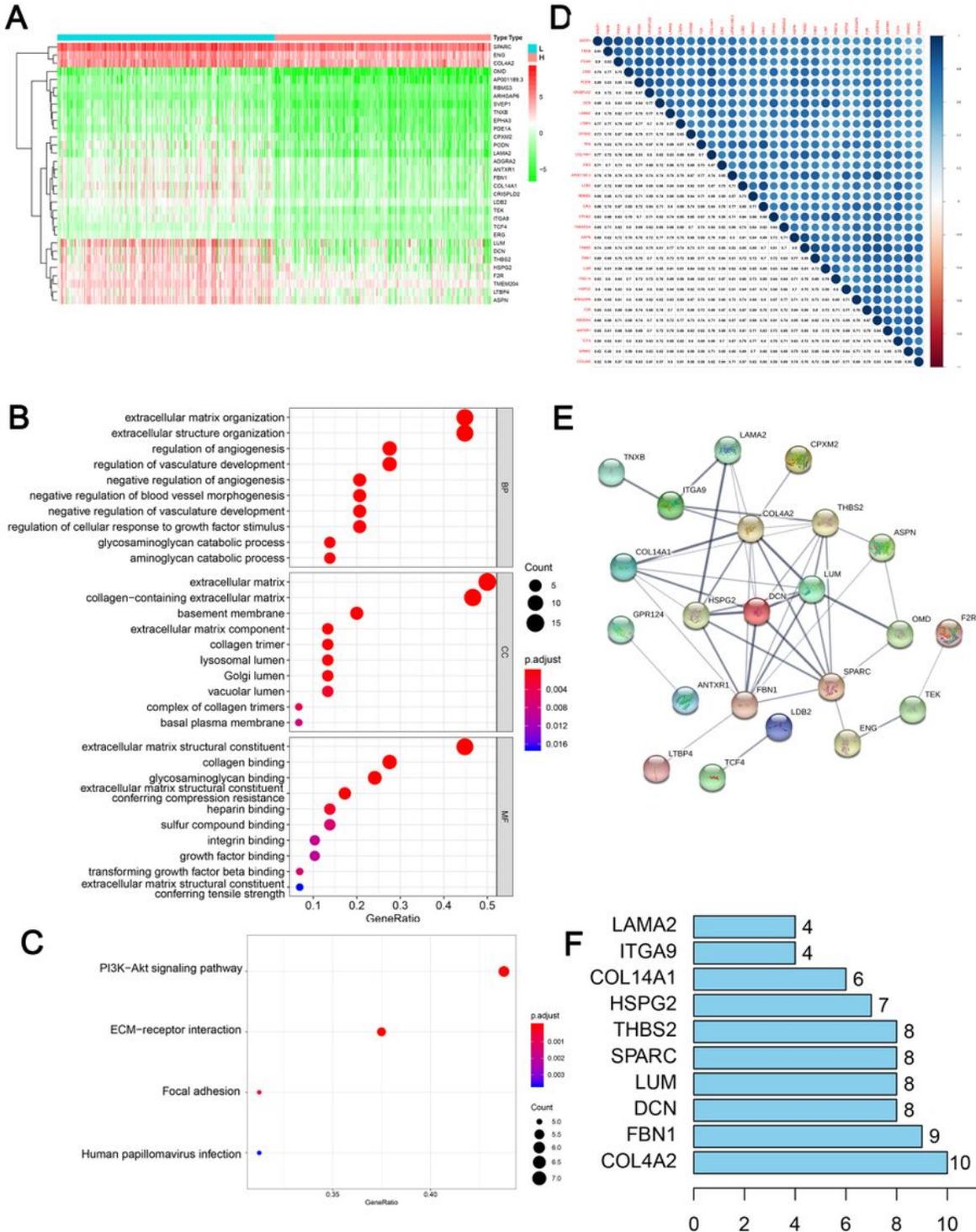


Figure 3

Functional analysis of yellow module genes (A). Heatmap of the expression level of yellow module genes. Samples were sorted into low mRNasi (L) and high mRNasi (H) groups. Red indicates high expression and green indicates low expression. Heatmap was drawn with $\log_2(\text{expression value}+0.01)$. (B). GO analysis of yellow module genes. (C). KEGG pathway analysis of yellow module genes. (D). Correlation between key genes at the transcriptional level. (E). PPI network analysis of yellow module genes. (F). edge number of yellow module genes in PPI network.

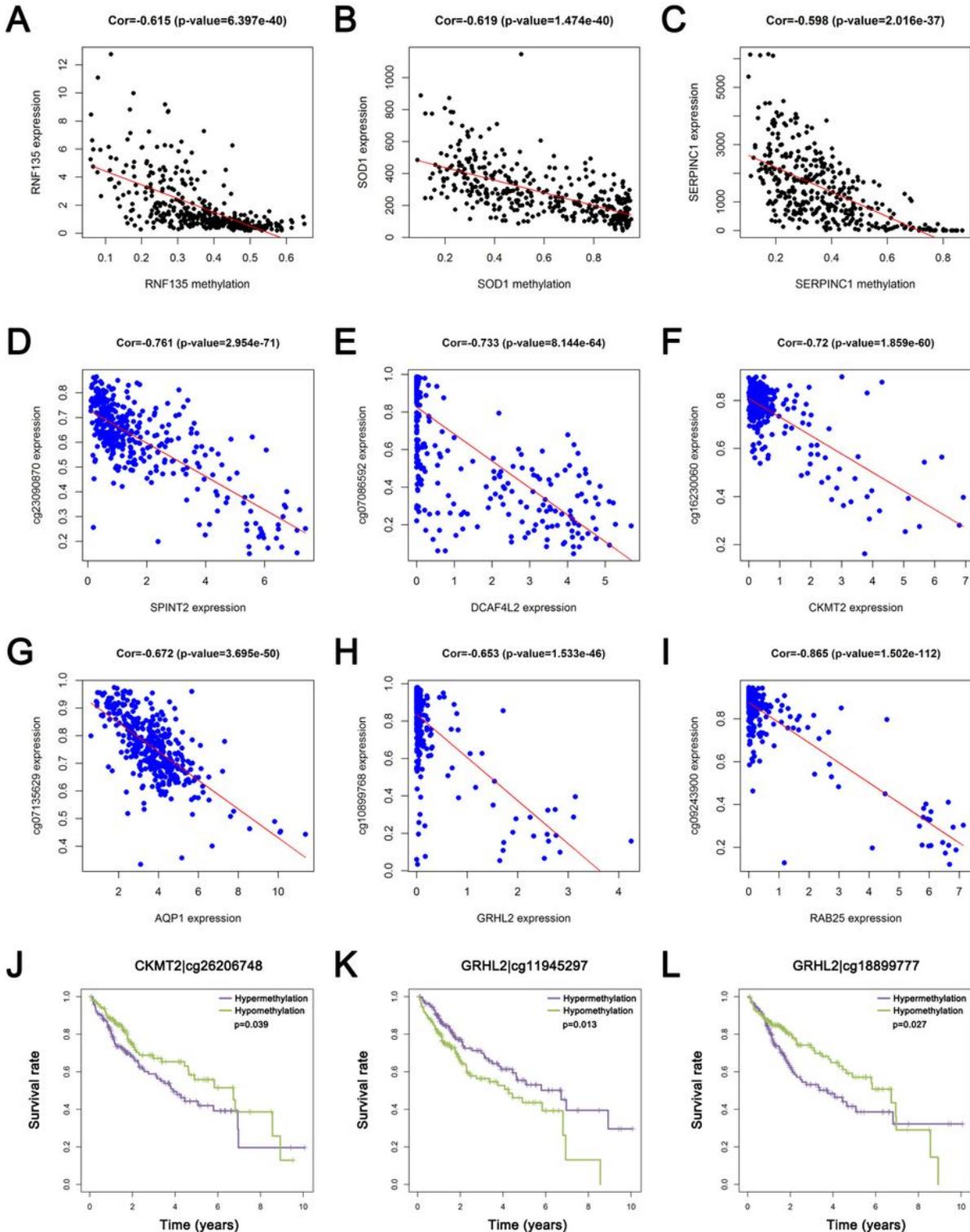


Figure 4

Screening of methylation-driven genes associated with stemness. (A-C). The representative methylation-driven genes. (D-I). Sites of methylation-driven genes associated with stemness. (J-L). Kaplan–Meier survival curves of the methylation site, including CKMT2|cg26206748, GRHL2|cg11945297 and GRHL2|cg18899777.

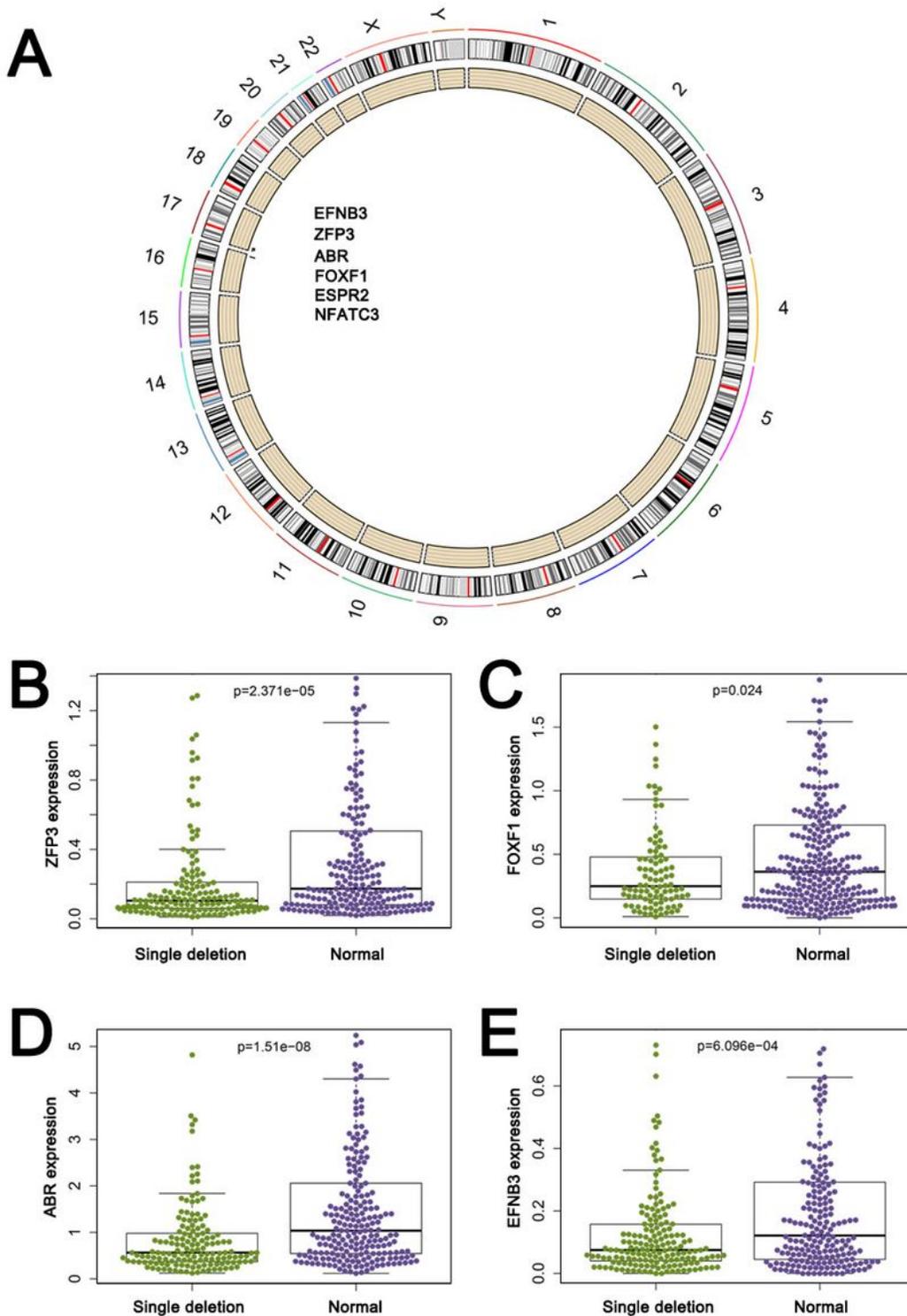


Figure 5

Screening of CNV-driven genes associated with stemness. (A). The location of CNV-driven genes on the chromosomes. (B-E). Visualization of CNV-driven genes associated with stemness.

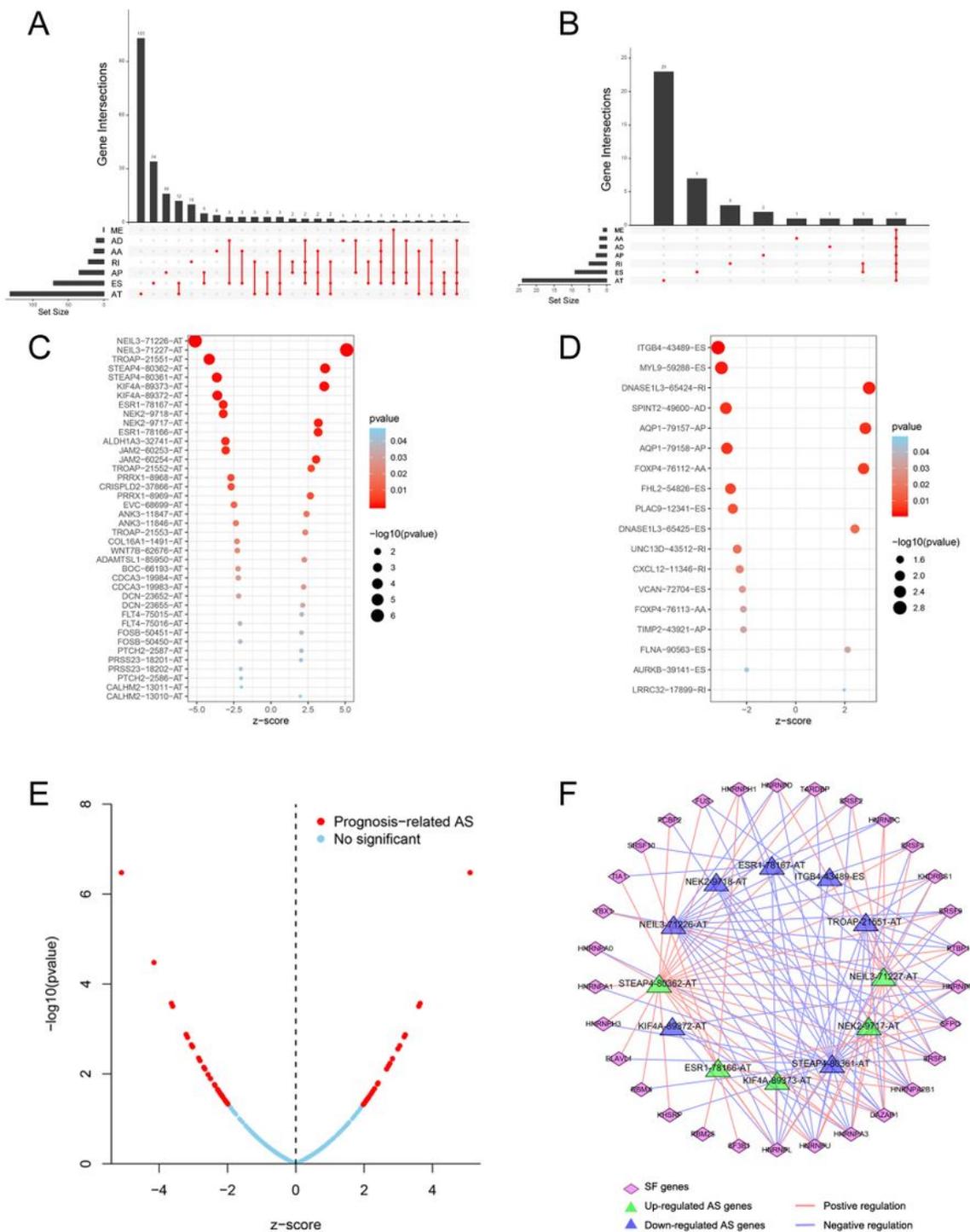


Figure 6

Construction of splicing correlation network associated with stemness. (A). UpSet plot of alternative splicing events in stemness-related genes. (B). UpSet plot of survival-related alternative splicing events in stemness-related genes. (C-D). Bubble plots of survival-related alternative splicing events in stemness-

related genes. (E). Volcano plots of alternative splicing events in stemness-related genes. (F). Splicing correlation network associated with stemness.

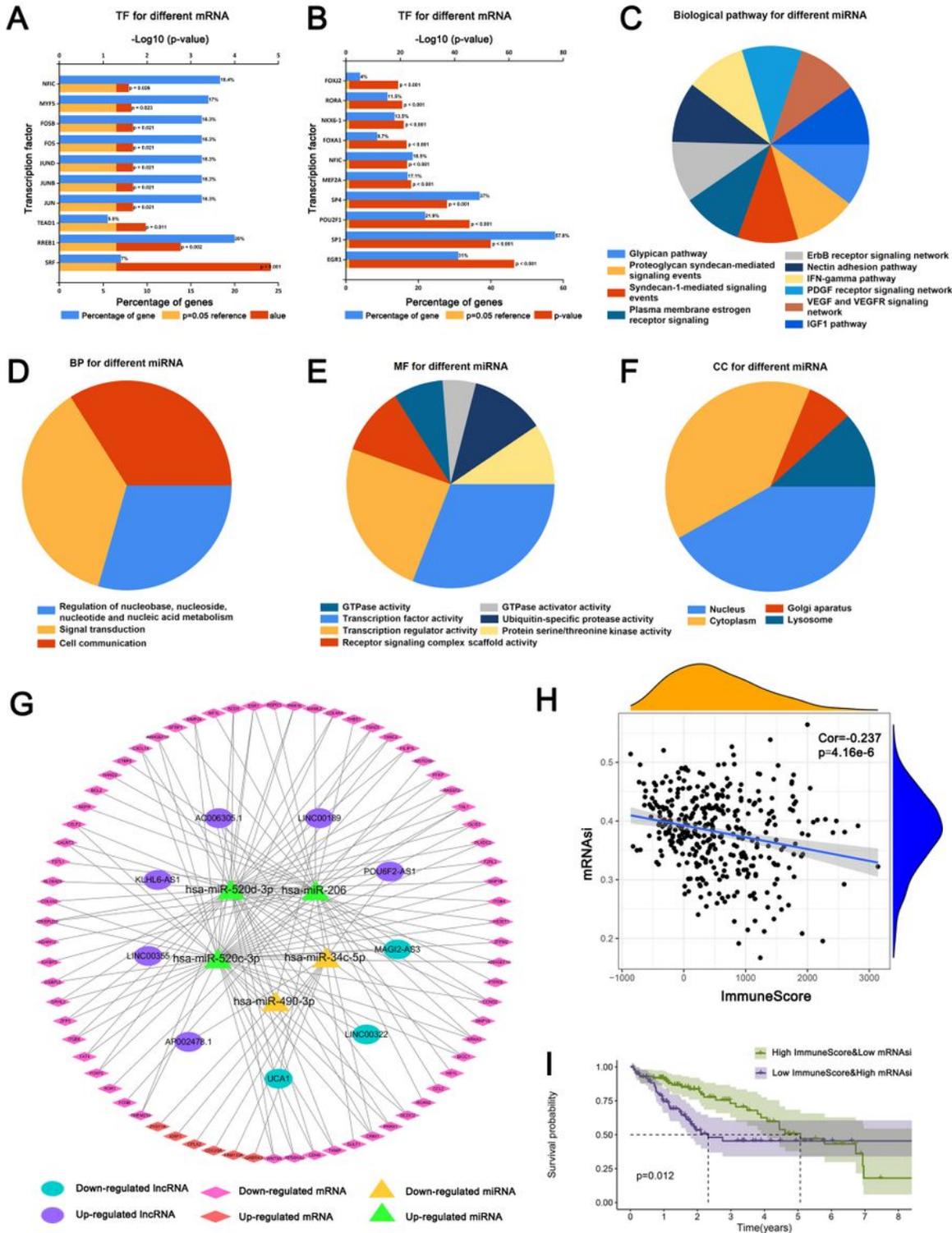


Figure 7

Construction of ceRNA network associated with stemness, and combined analysis of tumor immune microenvironment and mRNAsi in LC (A). Enrichment analysis of transcription factors for differentially expressed mRNAs was performed by FunRich. (B). Enrichment analysis of transcription factors for

differentially expressed miRNAs was performed by FunRich. (C). KEGG pathway analysis of differentially expressed miRNAs. (D-F). GO analysis of differentially expressed miRNAs. (G). The ceRNA network constructed by Cytoscape. (H). The correlation picture between immune scores and mRNAsi scores in LC. (I). Survival curves of mRNAsi scores combined with immune scores in LC.

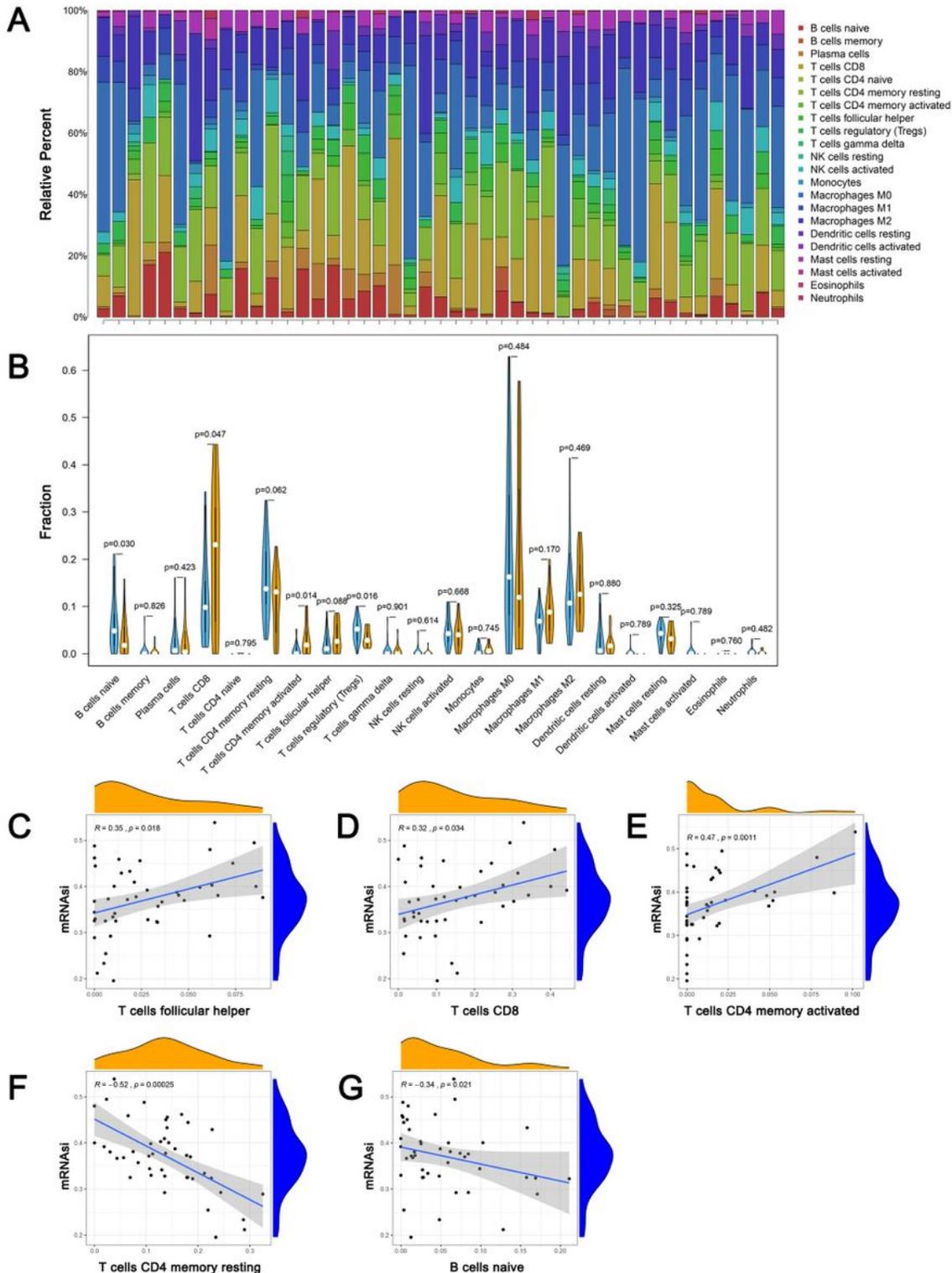


Figure 8

Immune cell composition in LC. (A). composition of infiltrating immune cells in LC patients. (B). Vioplot visualizing the differentially infiltrated immune cells in the low and high groups. (C-G). The correlation picture between infiltrating immune cells and mRANsi scores in LC.

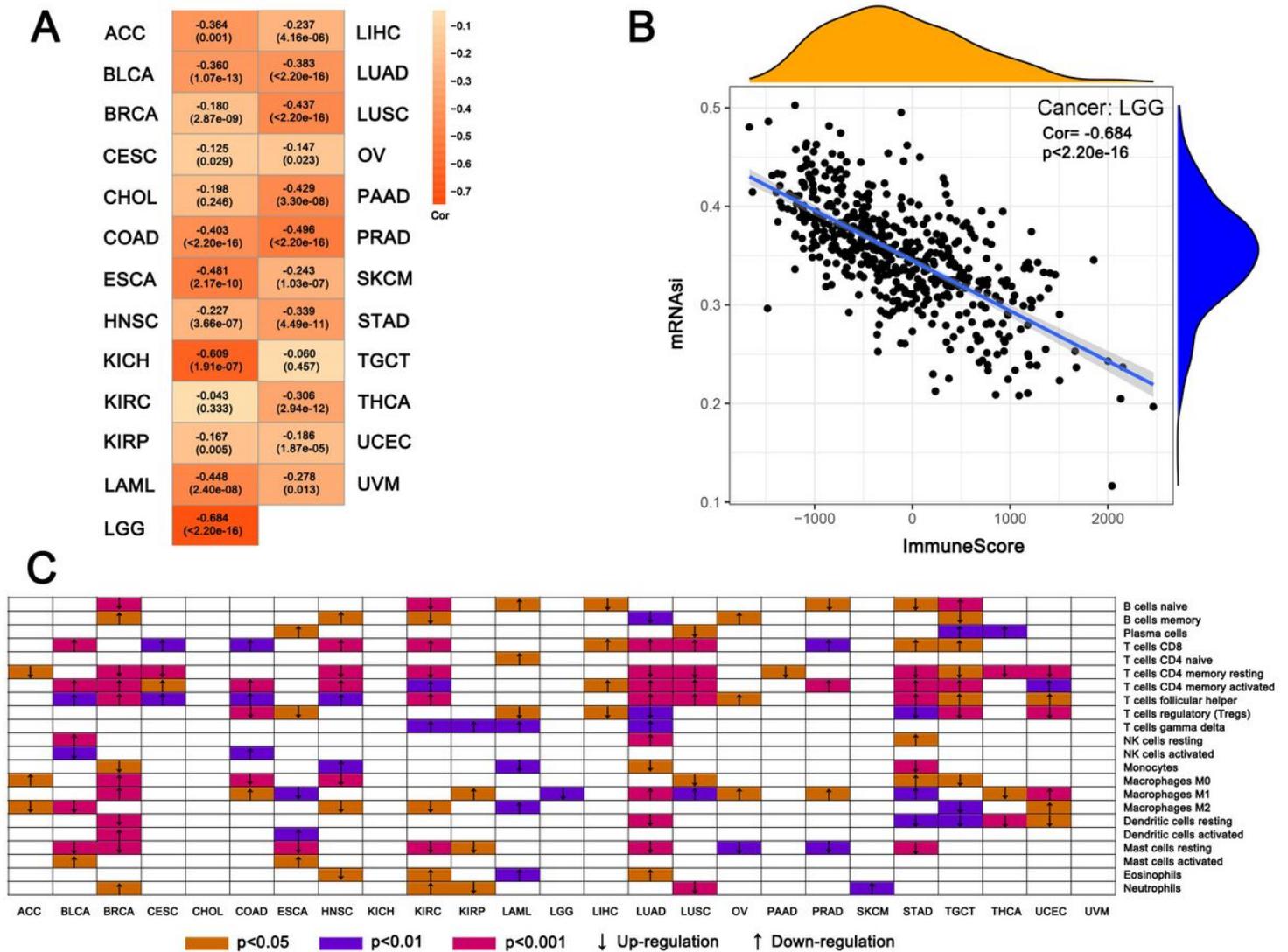


Figure 9

Analysis of tumor immune microenvironment and mRANsi in 25 types of tumors (A). The correlation coefficient between immune scores and mRANsi scores in 25 types of tumors. (B). The correlation picture between immune scores and mRANsi scores in LGG. (C) Comparison of infiltrating cells between the low and high mRNasi tumor groups in 25 types of tumors ACC, Adrenocortical carcinoma; BLCA, Bladder urothelial carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; ESCA, Esophageal carcinoma; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney chromophobe; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain lower grade glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; OV, Ovarian serous cystadenocarcinoma; PAAD,

Pancreatic adenocarcinoma; PRAD, Prostate adenocarcinoma; SKCM, Skin cutaneous melanoma; STAD, Stomach adenocarcinoma; TGCT, Testicular germ cell tumors; THCA, Thyroid carcinoma; UCEC, Uterine corpus endometrial carcinoma; UVM, Uveal melanoma.

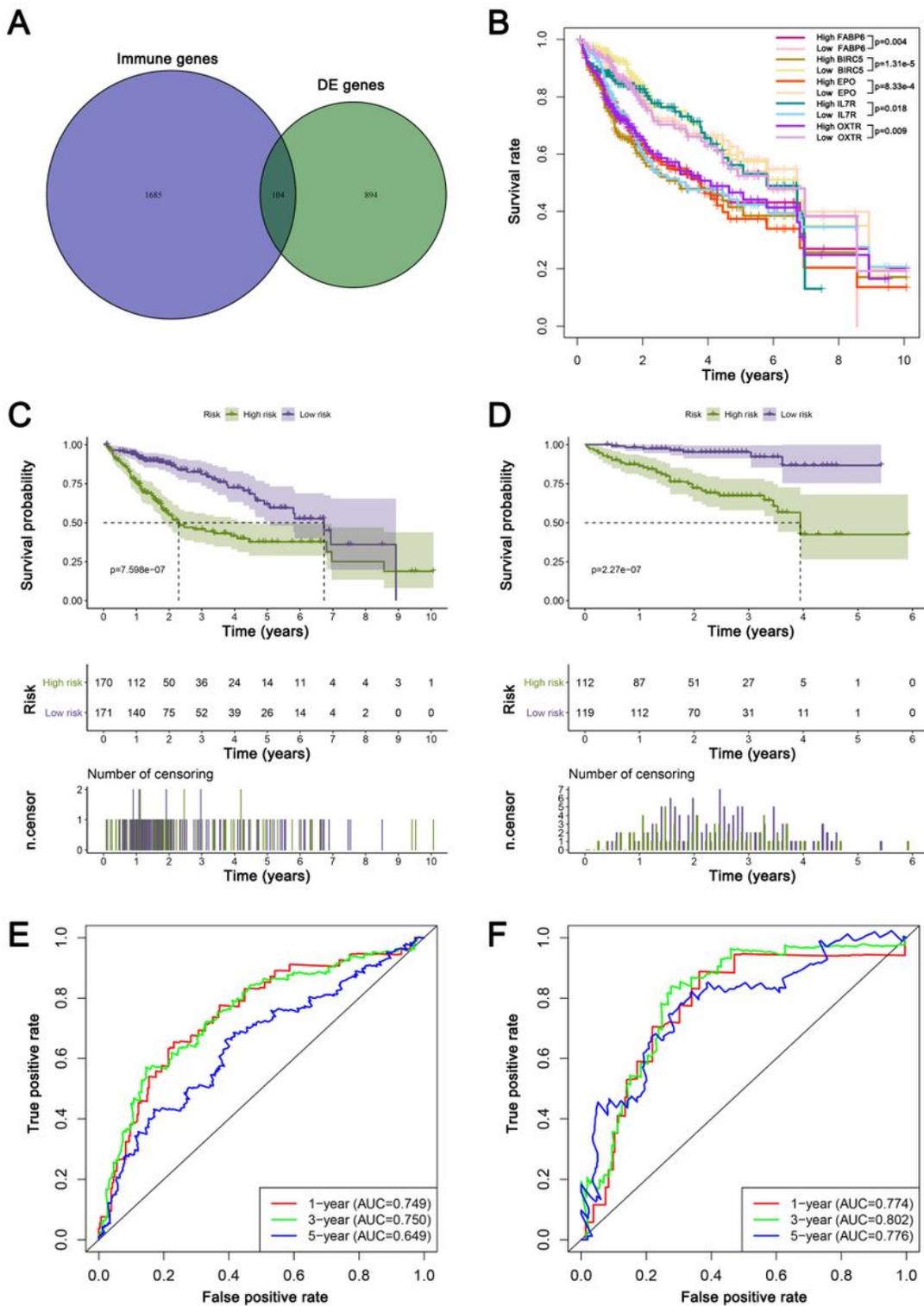


Figure 10

Construction and validation of immune-related prognostic model (A). The intersection of immune genes and DEGs between the low and high mRNasi tumor groups. (B). Kaplan–Meier survival curves of genes in

prognostic model. (C). Kaplan–Meier survival curves of prognostic model based on risk scores from TCGA patients. The median value at risk scores was defined as cut-of value. (D). Survival curves of risk scores from ICGC patients. The risk score of each patient was calculated using the same formula of prognostic model. The patients were divided into two groups using the same cut-of value, including low-risk group (119 samples) and high-risk group (112 samples). (E). ROC curve analysis of the prognostic predictors for LC patients from TCGA dataset. (F). ROC curve analysis of the prognostic predictors for LC patients from ICGC dataset.

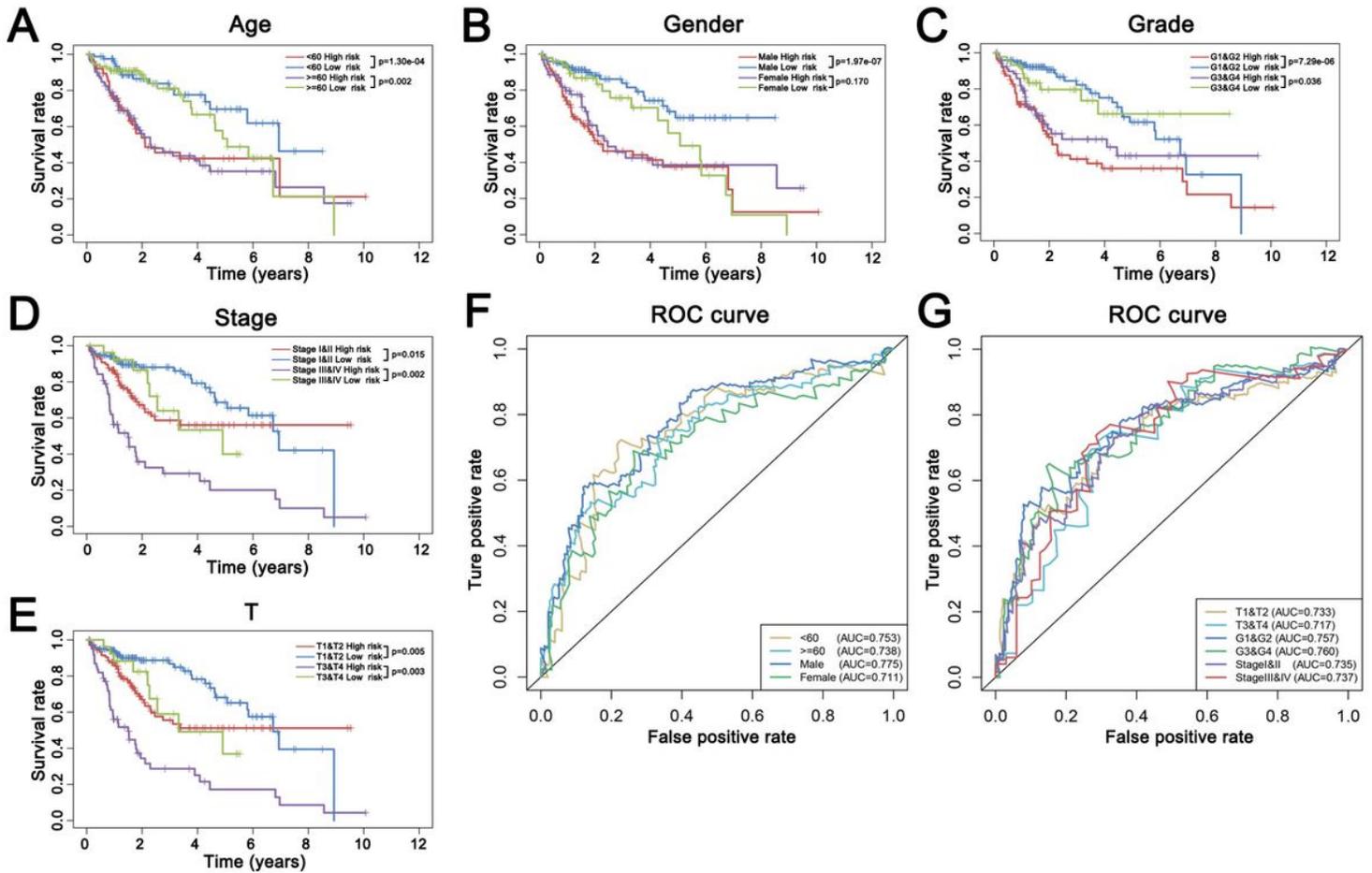


Figure 11

Survival analysis on the basis of the prognostic model stratified by clinicopathological factors. (A-E). Survival analysis on the basis of the prognostic model for patients who stratified by age, gender, grade, stage and tumor depths, respectively. The patients were divided into two groups using the above cut-of value. (F-G). ROC curve analysis of the prognostic predictors for LC patients who stratified by clinicopathological factors.

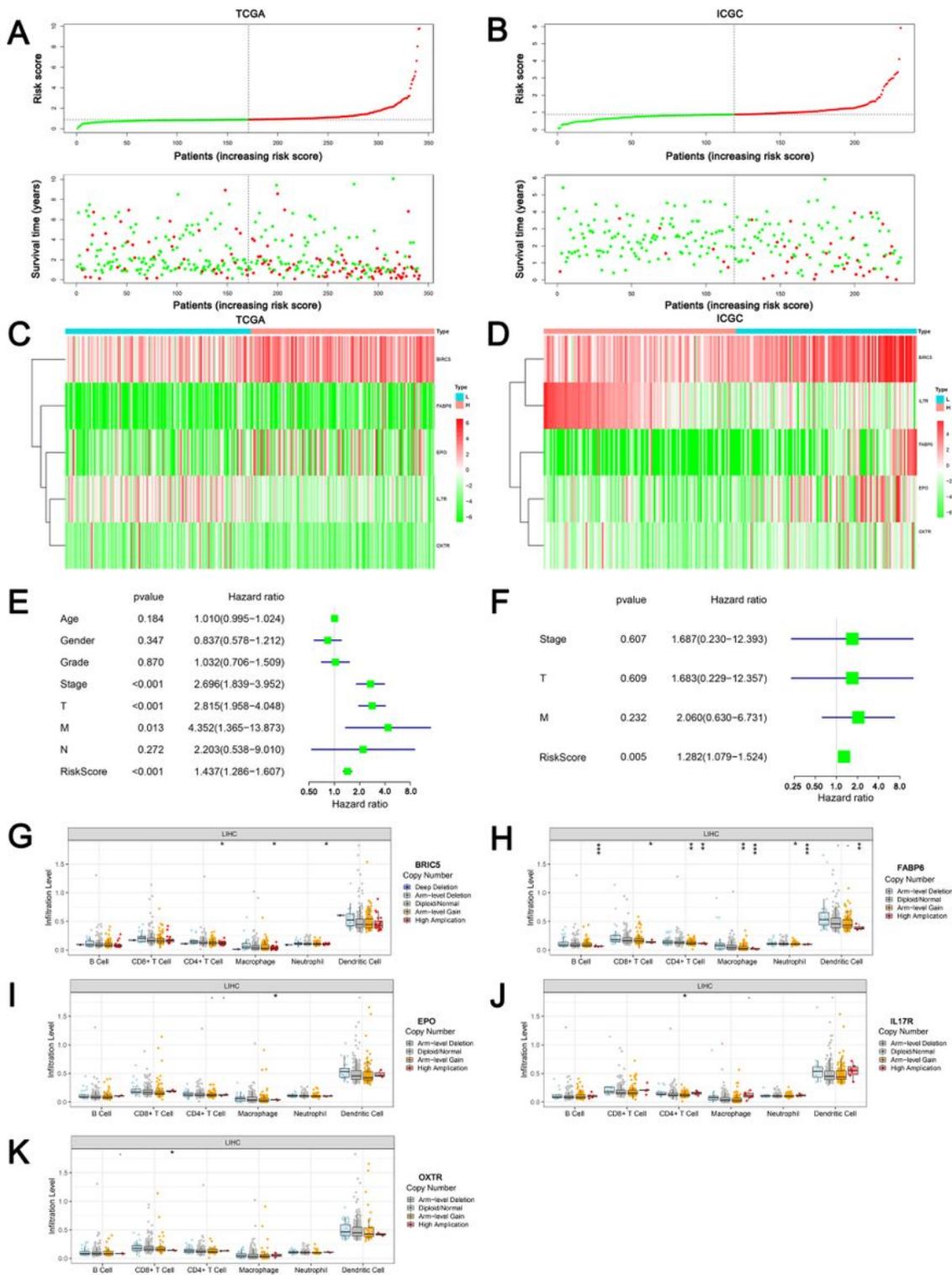


Figure 12

Distribution curve, COX analysis for risk score and relationship between prognosis model genes CNV and immune cells. (A-B). The top part displays the risk score's distribution curve and the bottom part shows and sorts the patients' survival state based on risk score. Patient data were from TCGA and ICGC, respectively. (C-D). The heatmap of the expression value of each prognostic gene. Data were from TCGA and ICGC dataset, respectively. Red indicates high expression and green indicates low expression.

Heatmap was drawn with $\log_2(\text{expression value} + 0.01)$. (E). Univariate COX analysis for risk score. Horizontal bars represent 95% CIs. (F). Multivariate COX analysis for risk score. Horizontal bars represent 95% CIs. (G-K) The relationship between immune cells and model genes CNV on basis of TIMER website (<https://cistrome.shinyapps.io/timer/>), including FABP6, BRIC5, IL17R, EPO and OXTR.

Supplementary Files

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

- [supplement1.tif](#)