Hypoxia associated multi-omics molecular landscape of tumor tissue in patients with hepatocellular carcinoma

Qiangnu Zhang[1,2,*], Lijun Qiao[1,2,*], Quan Liu[1,2], Xiangpan Kong[4], Jun Hu[5], Weibin Hu[1,3], Zongze Wu[1,3], Mingyue Li[1,3], Liping Liu[1,3]

[1]Department of Hepatobiliary and Pancreas Surgery, The Second Clinical Medical College, Jinan University, Shenzhen People's Hospital, Shenzhen 518020, Guangdong, China
[2]Integrated Chinese and Western Medicine Postdoctoral Research Station, Jinan University, Guangzhou 510632, Guangdong, China
[3]Department of Hepatobiliary and Pancreas Surgery, The First Affiliated Hospital, Southern University of Science and Technology, Shenzhen 518020, China
[4]Faculty of Science and Engineering, University of Groningen, Groningen 9747, Netherlands
[5]Institute of Pathology, University Clinic of Heidelberg, Heidelberg 69120, Germany
*Equal contribution

Correspondence to: Mingyue Li, Liping Liu; email: li.mingyue@szhospital.com; liu.liping@szhospital.com

ABSTRACT

The present study was designed to update the knowledge about hypoxia-related multi-omic molecular landscape in hepatocellular carcinoma (HCC) tissues. Large-size HCC datasets from multiple centers were collected. The hypoxia exposure of tumor tissue from patients in 10 HCC cohorts was estimated using a novel HCC-specific hypoxia score system constructed in our previous study. A comprehensive bioinformatical analysis was conducted to compare hypoxia-associated multi-omic molecular features in patients with a high hypoxia score to a low hypoxia score. We found that patients with different exposure to hypoxia differed significantly in transcriptomic, genomic, epigenomic, and proteomic alterations, including differences in mRNA, microRNA (miR), and long non-coding RNA (lncRNA) expression, differences in copy number alterations (CNAs), differences in DNA methylation levels, differences in RNA alternative splicing events, and differences in protein levels. HCC survival- associated molecular events were identified. The potential correlation between molecular features related to hypoxia has also been explored, and various networks have been constructed. We revealed a particularly comprehensive hypoxia-related molecular landscape in tumor tissues that provided novel evidence and perspectives to explain the role of hypoxia in HCC. Clinically, the data obtained from the present study may enable the development of individualized treatment or management strategies for HCC patients with different levels of hypoxia exposure.

INTRODUCTION

Hypoxia is a common microenvironmental feature of solid tumors including hepatocellular carcinoma (HCC). It leads to HCC cell proliferation activation, apoptosis inhibition, metabolic reorganization, immune escape, genetic instability, drug resistance, and angiogenesis [1, 2]. Hypoxia exposure of HCC tissues is related to the poor prognosis of patients [3]. Tumor tissue hypoxia caused by treatments such as transhepatic arterial chemotherapy and embolization may aggravate the malignant phenotype of HCC cells and impede the

therapeutic effect [4, 5]. Before formulating treatment measures for hypoxia, revealing the molecular mechanism is necessary. Studies have revealed the molecular mechanism by which hypoxia plays a role in tumors [6]. However, most of this evidence was obtained from *in vitro* cell studies and animal models, while little was obtained from the tissues of HCC patients because it is not convenient to evaluate the hypoxia exposure of the tissues of the patients [7]. The molecular landscape present in hypoxia-exposed HCC tissues still lacks a comprehensive description. In order to benefit patients, especially for the development of precision medicine and individualized medicine, a large amount of tissue-level evidence is urgently needed. Bhandari reported on molecular landmarks of tumor hypoxia across 19 cancer types [8]. Based on TCGA data, Ye et al revealed hypoxia-associated molecular features to aid hypoxia-targeted therapy [9]. These studies mentioned some data about HCC, but the signature they used to assess the degree of HCC tissue hypoxia was not HCC specific, and the description of the molecular feature caused by hypoxia in HCC was not comprehensive enough. In our previous works, a novel HCC-specific hypoxia signature containing 21 stable hypoxia-related genes was constructed using mRNA expression data. Based on the 21-gene signature, we grouped patients from 10 HCC cohorts into high hypoxia exposure and low hypoxia exposure groups. Next, we comprehensively compared changes in hypoxia-related molecular features in two groups from genomic, epigenomic, transcriptomic, and proteomic perspectives to deduce the hypoxia induced molecular landscape. We believe that the molecular landscapes revealed in the present study will provide useful information for developing therapy strategies of HCC.

RESULTS

Transcriptomic alterations in HCC patients with different hypoxia scores

In the previous studies, we established a hypoxia score system based on a novel HCC-specific 21-gene hypoxia signature that could be used to effectively estimate the hypoxia exposure in HCC tissues. In the 10 GEO datasets and the TCGA dataset, hypoxia scores of each tumor tissue samples were calculated. Patients were significantly grouped as high hypoxia exposure group and low hypoxia exposure group according to the hypoxia score (Supplementary Figure 1). First, we analyzed the differences in mRNA expression between HCC patients with high hypoxia scores (greater than the upper quartile) and those with low hypoxia scores (less than the lower quartile) cross the 11 HCC datasets. Here, mRNA with $\log_2 FC > 0.58$ or $\log_2 FC < -0.58$ and adjusted $P < 0.05$ were defined as differentially

expressed mRNA (DE-mRNA). In these cohorts, the proportion of DE-mRNA in the total mRNA measured was positively correlated with the interquartile range (IQR) of the liver cancer tissue hypoxia scores in this cohort (Figure 1A), suggesting that the difference in mRNAs was to some extent caused by the difference in hypoxia scores. We counted the frequency of each mRNA identified as a DE-mRNA in all cohorts and included DE-mRNAs with a frequency equal to or greater than 5 in the high frequency/DE-mRNA (HF/DE-mRNA) list. The mRNA changes in this list were relatively consistent among the 10 cohorts. A total of 371 mRNAs were selected, including 192 upregulated DE-mRNAs and 179 down-regulated DE-mRNAs. For a complete HF/DE-mRNA list, see Additional File 1: Supplementary Table 2. Based on the clinical data from the TCGA-LIHC and GSE14520 datasets, we performed survival analysis on 371 HF/DE-mRNAs (logrank test, cut-off = median expression level of candidate mRNA). Through Venny analysis, we integrated the survival data from the TCGA-LIHC and GSE14520 datasets and obtained 129 HF/DE-mRNAs related to HCC survival (logrank $P_{\text{TCGA-LIHC}}$ & logrank $P_{\text{GSE14520}} < 0.05$), which included 59 risk factors ($HR_{\text{TCGA-LIHC}}$ & $HR_{\text{GE14520}} > 1$) and 70 protective factors ($HR_{\text{TCGA-LIHC}}$ & $HR_{\text{GE14520}} < 1$). There were 51 HF/DE-mRNAs that met the logrank $P < 0.01$ and $HR < 0.7$ or > 1.3 requirements in both of TCGA-LIHC and GSE14520 (Figure 1B). The results of the HF/DE-mRNA survival analysis (logrank test) are shown in Additional File 1: Supplementary Table 3.

Hypoxia-inducible factor 1-alpha (HIF-1A) plays a core role in hypoxia. 2450 potential transcription targets genes of the HIF1A transcription factor predicted using the known transcription factor binding site motifs from the TRANSFAC Predicted Transcription Factor Targets database. 10%-15% of the upregulated DE-mRNAs in each cohort may be potential transcription targets of HIF-1A (Figure 1C). Among these potential HIF-1A transcription targets, DE-mRNAs with a frequency higher than 5 among ten cohorts are shown in Figure 1D; most are risk factors for survival. Besides, although HIF-1A protein levels are known to be regulated after translation under hypoxic conditions, 10 HCC datasets indicated that the mRNA level of HIF-1A significantly increased in the high hypoxia score group and showed a significantly positive correlation with the hypoxia score (Figure 1E, 1F).

To further reveal the functions of HF/DE-mRNAs, the enrichments of biological processes and pathways involving HF/DE-mRNA were analyzed using data from different sources, such as Gene Ontology (GO) biological processes, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Reactome Gene Sets, and Canonical Pathways (Figure 2A). In addition to the
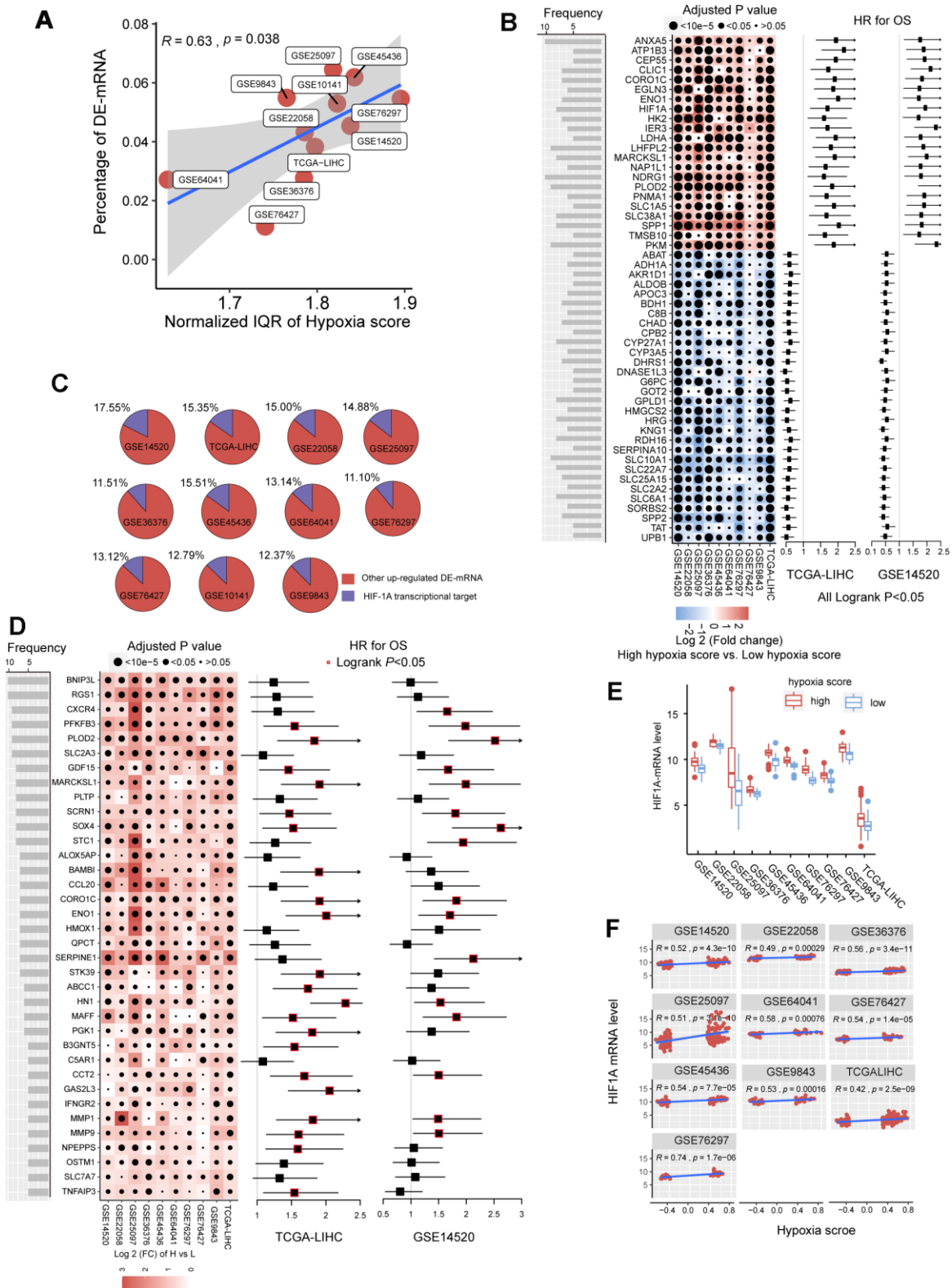
Figure 1. The mRNA alterations in hepatocellular carcinoma (HCC) patients with high hypoxia scores and low hypoxia scores. (**A**) Hypoxia scores were calculated based on the 21-gene hypoxia signature. According to the upper quartile and the lower quartile, patients were divided into a high hypoxia score group and a low hypoxia score group. In the 11 HCC cohorts, the percentage of differentially expressed (DE)-mRNAs among all mRNAs measured was positively proportional to the interquartile range (IQR) of the hypoxia scores. (**B**) A total of 51 high frequency/DE-mRNAs (HF/DE-mRNAs) are correlated to HCC patient survival in both TCGA-LIHC and GSE14520. The heat map shows the difference in the expression of these mRNAs between the high hypoxia score group and the low hypoxia score group in the 10 HCC

cohorts, that is, the log2 (fold change) between the two groups. The forest plot indicates the hazard ratios (HRs) of these mRNAs for OS in the survival analysis (all logrank P < 0.01, HR < 0.7 or > 1.3, cut-off value = median expression level). (**C**) The percentage of transcription targets with differentially expressed hypoxia-inducible factor 1-alpha (HIF-1A) in a dataset for all DE-mRNAs in the dataset. (**D**) Thirty-six mRNAs may function as transcription targets of HIF-1A, and the upregulation trends are consistent in the 10 HCC datasets. The heat map shows the difference in the expression of these mRNAs between the high hypoxia score group and the low hypoxia score group. The forest plot indicates the HRs of these mRNAs for OS in the survival analysis (cut-off = median expression level). (**E**) The differences in HIF-1A mRNA expression levels between the high hypoxia score group and the low hypoxia score group in 10 HCC datasets. (**F**) Correlations between HIF-1A mRNA expression levels and hypoxia scores for the 10 HCC datasets.

response to hypoxia, HF/DE-mRNAs are mainly involved in biological processes related to metabolism, including glucose metabolism, lipid metabolism, and amino acid metabolism. As expected, pathway enrichment analysis results for data from multiple sources all showed significant enrichment of the HIF-1 pathway and glucose metabolism pathways; other enriched pathways were mainly related to various metabolic pathways. Besides, some classical tumor-related pathways, such as the PI3K-AKT signaling pathway and activated protein kinase (AMPK) signaling pathway, were also significantly enriched (Additional File 1: Supplementary Table 4). The pathways related to the regulation of the extracellular matrix were also associated with HF/DE-mRNAs. Therefore, we hypothesized that hypoxia exposure might affect the extracellular matrix, which determines tumor invasion and metastasis. To find the connection between terms in the enrichment analysis of biological processes and pathways, we clustered the terms and constructed a network (Figure 2B). The name of each cluster was the name of the most representative term, and the node size was the number of genes in the term. The names of all nodes can be found in Additional File 2. For the HF/DE-mRNA translation products, protein-protein interaction (PPI) enrichment analysis was carried out with the following databases: BioGrid6, InWeb_IM7, and OmniPath8. We constructed a PPI enrichment network of physical interactions using molecular complex detection (MCODE) (Figure 2C). The names of all nodes are listed in Additional File 3. The annotation of each MCODE is shown in Additional File 1: Supplementary Table 5. In addition to focusing on HF/DE-mRNAs, we used gene set enrichment analysis (GSEA) to reveal differences between hallmark gene sets between the high hypoxia score group and the low hypoxia score group based on all mRNA differences between the two groups (Figure 2D). The hypoxia gene set was upregulated in all cohorts with high hypoxia scores. The genes composing the glycolysis gene set was upregulated in the high hypoxia score groups of 9 cohorts. Other upregulated gene sets with high consistency in the 10 cohorts included the P53 pathway, PI3K/AKT/mammalian target of rapamycin (mTOR) signaling, TNFA signaling via NFKB, unfolded protein response, TGF-beta signaling, and MTORC1 signaling. Although the fatty acid metabolism pathway was

downregulated in the high hypoxia score groups in the TCGA-LIHC and GSE14520 cohorts, downregulation was not consistent among the other pathways. After applying a bimodality filter and weighted gene correlation network-based clustering, Hoadley's team identified 22 nonredundant gene programs related to the biological behaviors of tumors [10]. We found 15 gene programs that were significantly different between the high hypoxia score group and the low hypoxia score group (Figure 2E), and the single-sample GSEA z-scores for some cancer-promoting gene programs increased in patients with high hypoxia scores.

We used TCGA-LIHC data to analyze the differences in microRNAs (miRs) in HCC tissues with high hypoxia scores and low hypoxia scores. The miRs with $\log_2 FC > 0.58$ or $\log_2 FC < -0.58$ and adjusted $P < 0.05$ were defined as differentially expressed miRs (DE-miRs). We found a total of 63 DE-miRs, including 39 upregulated DE-miRs and 24 downregulated DE-miRs. Survival analysis showed that some DE-miRs were related to the OS rate (logrank test, cut-off = median expression level) of HCC patients (Figure 3A and Additional File 1: Supplementary Table 6). In our study, miR-210-3p had the smallest adjusted *P value* among the upregulated DE-miRs ($\log_2 FC = 2.41$). Survival analysis using the median expression level as the cut-off showed that miR-210-3p had the largest HR (logrank $P < 0.05$) and that high miR-210-3p expression indicated a poor prognosis. Additionally, the passenger strand of miR-210 (miR-210-5p) was significantly upregulated in HCC tissues with high hypoxia scores, and high miR-210-5p expression also indicated a poor prognosis. Among the downregulated DE-miRs, miR-139-5p had the smallest adjusted *P value* ($\log_2 FC = -0.68$). The response of miR-139-5p to hypoxia has not been reported. We found that low miR-139-5p expression indicates poor outcomes of HCC patients. Next, we predicted target mRNAs of DE-miRs. Combined with the HF/DE-mRNAs list, we obtained 2 independent DE-miRs-HF/DE-mRNAs networks, including upregulated DE-miRs/downregulated DE-mRNAs and downregulated DE-miRs/upregulated DE-mRNAs (Supplementary Figure 2). HIF-1A mRNA has a targeted relationship with some downregulated miRs, such as miR-101-3p and miR-194-3p. These miRs can explain the increase in HIF-1A mRNA expression in the high hypoxia score groups of the 10 cohorts.
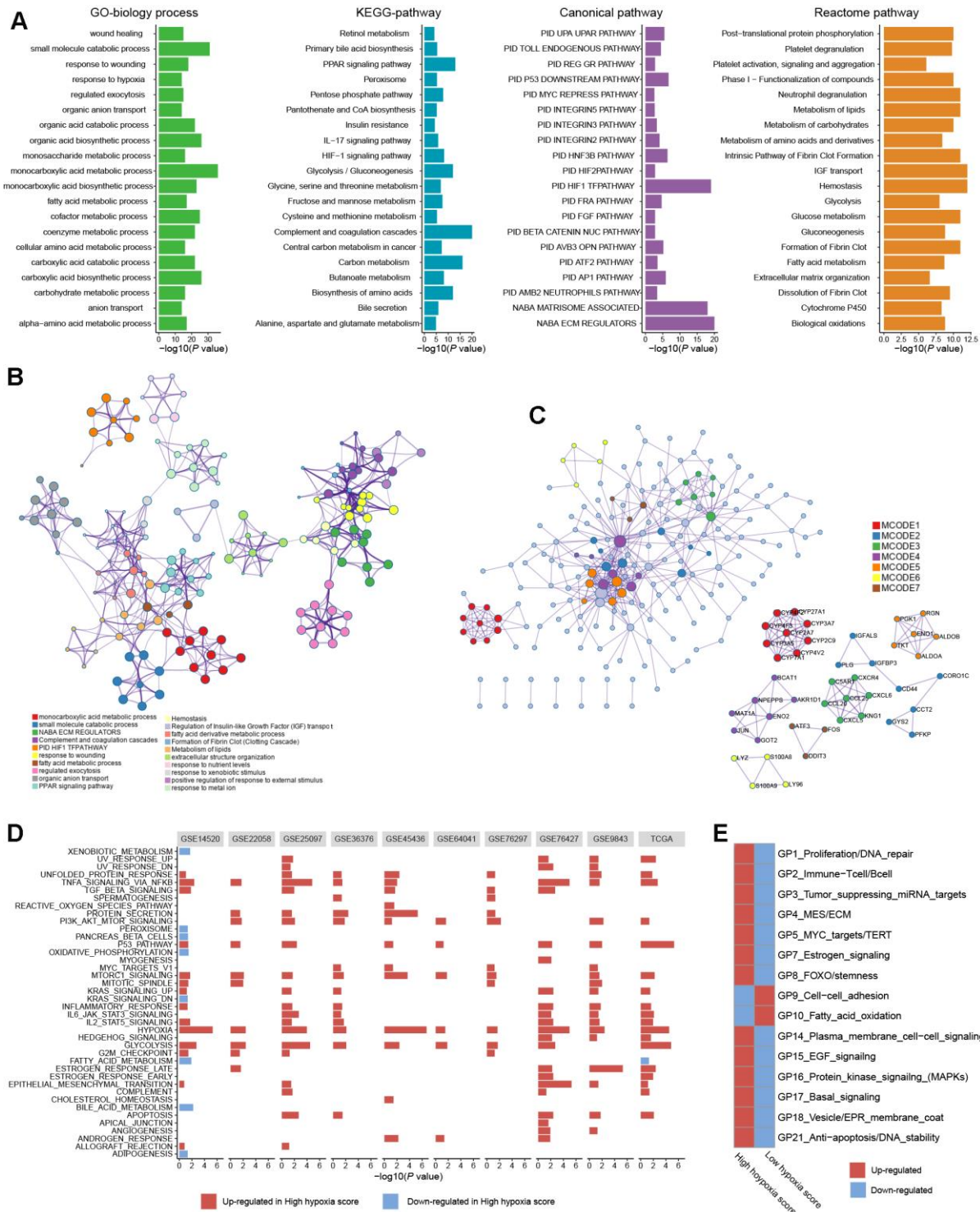
Figure 2. Biological processes and pathway functional enrichment analysis of 371 high frequency/DE-mRNAs (HF/DE-mRNAs) extracted from 10 hepatocellular carcinomas (HCC) datasets. (**A**) The top 20 (sorted by P-value) from each enrichment analysis result for GO biological processes of HF/DE-mRNAs and the pathway enrichment analysis results for 3 data sources. (**B**) A clustering network formed by correlated terms from the functional enrichment analysis. (**C**) Protein-protein interaction (PPI) enrichment network of HF/DE-mRNA translation products constructed based on the molecular complex detection (MCODE) algorithm. (**D**) Gene set enrichment analysis (GSEA) of HF/DE-mRNAs in 10 datasets, showing the pathways with P < 0.05 and false detection rate (FDR) <0.25. The reference gene sets are the hallmark gene sets. In the TCGA-LIHC dataset, (**E**) the differences between the high hypoxia score group and the low hypoxia score group in 15 gene programs related to biological behaviors of tumors. These gene programs were identified by Hoadley's team, and the upregulation and downregulation trends were calculated as single-sample GSEA z-scores.

A survival-related refined network was obtained by combining the survival analysis results (Figure 3B). In this refined DE-miR/DE-mRNA network, all nodes were associated with the survival rate of HCC patients (logrank $P < 0.05$), and a negative correlation between nodes was indicated ($r < -0.4$ and $P < 0.05$). For example, miR-194-5p showed low expression in the high hypoxia score group, which was a protective factor for the survival rate. Its potential target genes, SOX4, HK2, MARCKS, and LHFPL2, showed high expression in the high hypoxia score group and were risk factors for the OS rate. The expression of miR-194-5p was significantly negatively correlated with the mRNA expression of SOX4, HK2, MARCKS, and LHFPL2. The elements in the refined DE-miR/DE-mRNA network should receive more attention. We performed KEGG pathway enrichment analysis on all target genes of DE-miRs. The enrichment results of the top 20 (according to the P-value) were shown in Figure 3C. Some important classical pathways related to tumor development, such as the Hippo pathway, TGF-beta pathway, Ras signaling pathway, mTOR pathway, PI3K-AKT pathway, Wnt pathway, and AMPK pathway, were involved. This result further suggests that these DE-miRs may have an unignorable role in HCC tissues with high hypoxia scores. Besides, more than 80% (54/63) of the DE-miRs had at least one target gene enriched in the HIF-1 signaling pathway. This confirms that DE-miRs identified by the 21-gene hypoxia signature is indeed hypoxic-related.

Through the 21-gene signature, we revealed the presence of long non-coding RNAs (lncRNAs) that respond to and influence hypoxia exposure. In HCC tissues of patients with high hypoxia scores and low hypoxia scores, we found 719 differentially expressed lncRNAs (DE-lncRNAs), including 499 upregulated DE-lncRNAs and 220 downregulated DE-lncRNAs. The top-50 DE-lncRNAs (sorted by adjusted P) and survival analysis results (logrank test, cut-off = median expression level) are shown in Figure 3D. The complete lists of DE-lncRNAs and survival analysis results can be found in Additional File 1: Supplementary Tables 7, 8, respectively. miR210HG exhibited the most significant change ($\log_2 FC = 2.21$) among the upregulated DE-lncRNAs. Higher miR210HG suggested poorer survival ($HR = 1.82$, logrank $P < 0.05$). Similarly, lncRNAs AC124798 and AC061992 were also upregulated in the high hypoxia score group and suggested poor prognosis. In addition, lncRNAs LNC00671 and FAM99A were downregulated in the high hypoxia score group, which might be protective factors for prognosis. Based on competing endogenous RNA (ceRNA) theory and expression changes, we constructed a DE-lncRNA-DE-microRNA-HF/DE-mRNA network (Figure 3E). The relationship between each node and the survival rate for

HCC patients was identified, and a negative correlation between nodes was also indicated. The miR-lncRNA relationship in the network was supported by experimental evidence (provided by DIANA-LncBase). Taking lncRNA-SNHG12 as an example, it exhibits high expression in the high hypoxia score group and is a risk factor for the survival of HCC patients. lncRNA-SNHG12 and miR-194-3p may have sequence complementarity. miR-194-3p was low in the high hypoxia score group and thus is a protective factor for HCC patient survival. The target mRNAs of miR-194-3p were TM4SF1 and HIF-1A. TM4SF1 and HIF-1A showed high expression in the high hypoxia score group and thus were risk factors for HCC patient survival. TM4SF1, HIF-1A, and SNHG12 showed a significantly negative correlation with miR-194-3p. Therefore, a hypoxia-responsive lncRNA-SNHG12/miR-194-3p/TM4SF1 or HIF-1A ceRNA network is likely to present in the cancer tissues of HCC patients, and the ceRNA network is involved in tumor development and is related to patient prognosis. Leaving out sequence complementarity, we constructed the co-expression networks of all HF/DE-mRNAs, DE-miRs, and DE-lncRNAs ($|Pearson\ r| > 0.8$ and $P < 0.05$), including a positive co-network (Supplementary Figure 3A, Pearson $r > 0.8$ for all nodes) and a negative co-expression network (Supplementary Figure 3B, Pearson $r < -0.8$ for all nodes); the hub genes in the 2 networks were SERPINC1 and PKM, respectively. SERPINC1 was significantly downregulated in the high hypoxia score groups of 6 cohorts while PKM was significantly upregulated in the high hypoxia score groups of 10 cohorts (Supplementary Figure 3C). Compared with those in normal tissues, SERPINC1 was significantly lower and PKM was significantly higher in HCC tissues (Supplementary Figure 3D). Combined with the survival analysis results (Supplementary Figure 3E), we speculated that SERPINC1 and PKM play important roles, namely, cancer-suppressing and cancer-promoting functions, respectively, under hypoxia exposure. Supplementary Figure 4 displays all mRNAs and lncRNAs related to HIF-1A mRNA levels ($|Pearson\ r| > 0.4$ and $P < 0.05$), and the relationship between nodes and survival was indicated. Among miRs, only microRNA-194-3p and microRNA-194-5p had a significantly negative correlation with HIF-1A. These data may help to explain the differences in HIF-1A mRNA between groups with high hypoxia scores and reflect the core role of HIF-1A mRNA.

Genomic alterations in HCC patients with different hypoxia scores

TCGA-LIHC data were used to reveal somatic copy number aberrations (CNAs) and somatic single-nucleotide variants (SNVs) in HCC patients with
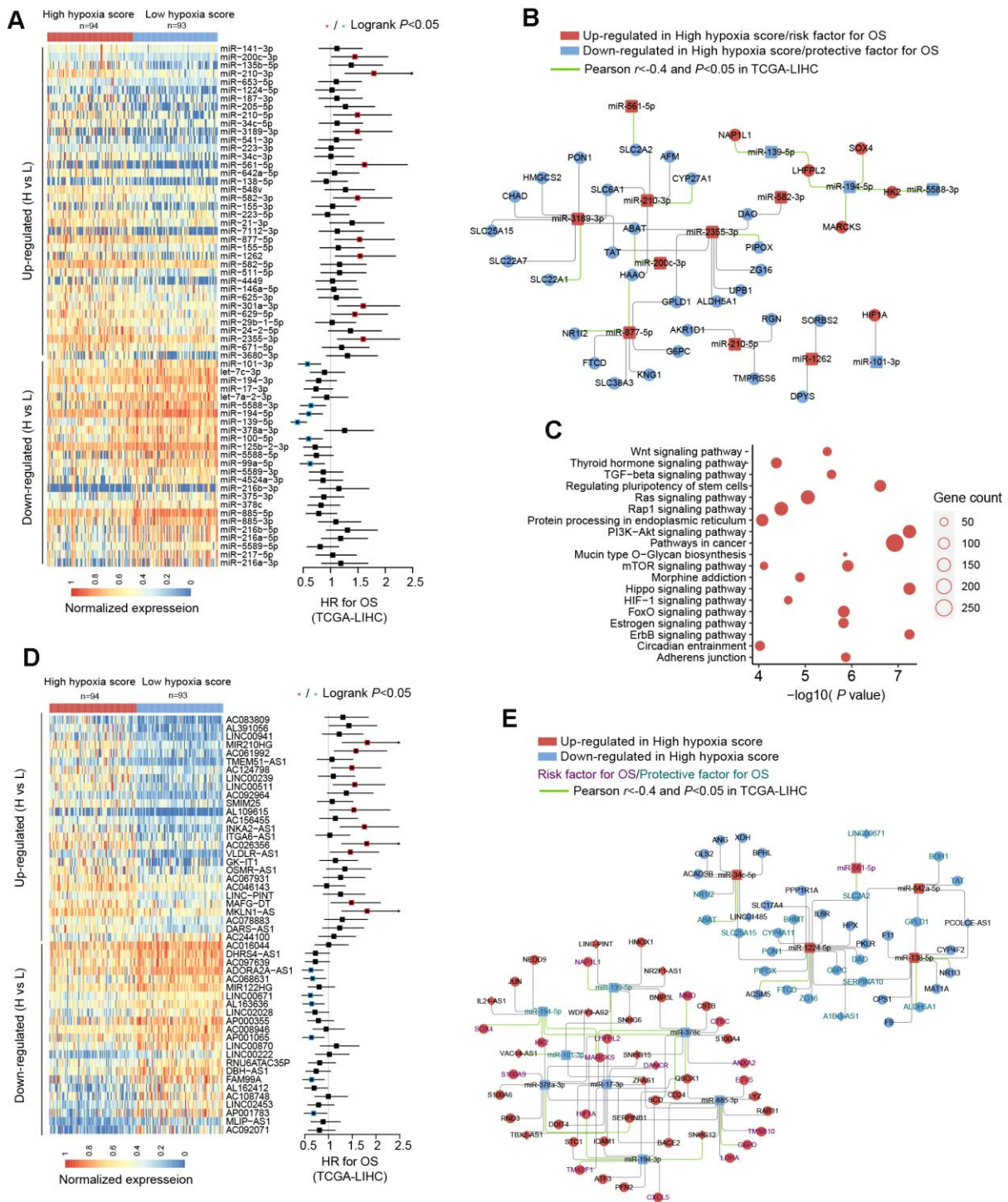
Figure 3. miRNA and long non-coding RNA (lncRNA) alterations in HCC patients with high hypoxia scores and low hypoxia scores. (**A**) A total of 63 DE-miRNAs were significantly upregulated or downregulated in the high hypoxia score group. The forest plot indicates the hazard ratios (HRs) of these miRNAs for overall survival (OS) in the survival analysis (logrank test). (**B**) Some DE-miRNAs and HF/DE-mRNAs constitute a survival-related target interaction network. All nodes in the network are correlated with HCC patient survival in TCGA-LIHC (logrank P < 0.05, cut-off = median expression level). The correlations between the nodes were calculated using Pearson correlation analysis. (**C**) Top 20 (sorted by P-value) KEGG pathway enrichment analysis results from 633 DE-miRNA target genes. (**D**) Top 50 (sorted by adjusted P-value) DE-lncRNAs that were significantly upregulated or downregulated in the high hypoxia score group. The forest plot indicates the hazard ratios (HRs) of these lncRNAs for OS in the survival analysis (logrank test). (**E**) The refined DE-lncRNA–DE-miRNA–HF/DE-mRNA ceRNA network. The correlations between nodes were calculated by Pearson correlation analysis. The survival data were from TCGA-LIHC. The cut-off is the median expression level.

different hypoxia scores. First, from the overall differences in gene-level CNAs in patients with high hypoxia scores (greater than the upper quartile) and low hypoxia scores (less than the lower quartile), we found that CNAs of approximately 13.7% (3396/24769) of the genes were concentrated in the high hypoxia score group (Supplementary Figure 5A and Additional File 1: Supplementary Table 9). CNA events in 71 cancer genes (according to the definition of the Precision Oncology Knowledge Base (OncoKB) cancer gene list) were significantly differentially distributed between the two groups (Figure 4A). The copy number gain to copy number loss ratios for most genes were significantly increased in the high hypoxia score group. The roles and CNA tendencies (according to the proportions of homozygous deletions, single copy deletions, low-level copy number amplification, and high-level copy number amplification) of these 71 cancer genes were provided.

For example, CDK4 is an oncogene, and its CNAs in the high hypoxia score group are mostly copy number gains. IRF1 is a tumor suppressor gene, and its CNAs in the high hypoxia score group were mostly copy number losses. The occurrence frequency of CNAs of some cancer genes was high in patients in the TCGA-LIHC cohort. Next, we analyzed the difference in gene-level SNVs between the two groups of patients. Unfortunately, we did not obtain much evidence that indicated a strong connection between SNVs and hypoxia score. Only 4 genes were significantly different in the incidence of SNVs between the two groups (Figure 4B). Among them, the proportion of non-silent mutations only increased in ADAMTS19 in the high hypoxia score group. Besides, we found that SNVs in 172 genes tended to be concentrated in the low hypoxia score group or the high hypoxia score group (Supplementary Figure 5B). However, because the overall mutation frequency of these genes was not high, there was no statistically significant difference in the distribution of SNVs between the two groups.

Epigenetic alterations in HCC patients with different hypoxia scores

We found that there were significant differences in the methylation levels at 464 gene loci between patients with high hypoxia scores and patients with low hypoxia scores and that the methylation levels at most loci were significantly reduced in the high hypoxia score group (Supplementary Figure 5C and Additional File 1: Supplementary Table 10). The methylation level increased at only a few loci in the high hypoxia score group. We jointly analyzed methylation levels and mRNA expression levels and found a significant increase in mRNA expression levels of 30 genes in the high hypoxia score group (data from TCGA-LIHC) and

a simultaneous decrease in their methylation levels (Figure 4C). The mRNA expression levels of these genes showed a significantly negative correlation with their methylation levels, and the high mRNA expression of most of these genes is a risk factor for HCC patient survival. The methylation loci corresponding to these 30 genes and their distance from the transcription start site (TSS) are shown in Additional File 1: Supplementary Table 11. Notably, the degree of methylation of HIF-1A was also reduced in the high hypoxia score group, which provides another explanation for the increase in HIF-1A mRNA expression in the high hypoxia score group. From the comparison of DNA methylation levels of the 30 genes in normal tissues and HCC tissues, the DNA methylation levels of some genes in normal tissues were not different from those in HCC tissues with low hypoxia scores but were higher than those in HCC tissues with high hypoxia scores (such as RGS2), and the DNA methylation levels of some other genes were highest in normal tissues, followed by those in HCC tissues with low hypoxia scores, and lowest in HCC tissues with high hypoxia scores (such as CLIC1).

Chromosome instability occurs throughout the development and progression of tumors. The increase in chromosome instability will lead to HCC cell growth and enhanced invasiveness [11]. The 70-gene chromosome instability (CIN70) signature constructed by Carter et al. can effectively assess the chromosome instability of tumor cells [12]. We compared the CIN70 scores between patients with high hypoxia scores and patients with low hypoxia scores. The data showed that in 9 datasets, the CIN70 scores were significantly increased in the high hypoxia score groups (Figure 4D), suggesting that hypoxia exposure may cause chromosome instability in HCC cells.

Alternative splicing (AS) is an epigenetic feature that plays an important role in tumorigenesis and development [13]. Based on the TCGA-LIHC data, we found that the occurrences of 713 AS events (involving 681 mRNAs) differed significantly between the two groups and that most of these events were alternative terminator (AT), exon skipping (ES), and alternative promoter (AP) events (Figure 4E). In the high hypoxia score group, the percent spliced in (PSI) value of retained intron (RI) events of EPHB2 mRNA had the maximum reduction (adjusted P value < 0.001), and the AT events of GULP1 mRNA had the most increased PSI (adjusted P value < 0.01). The Cox analysis results suggested that the AS events with differences in occurrence between the two groups were related to HCC patient survival. Detailed information on these differentials AS events is provided in Additional File 1: Supplementary Table 12. The 681 mRNAs involved in differential AS events were subjected to functional
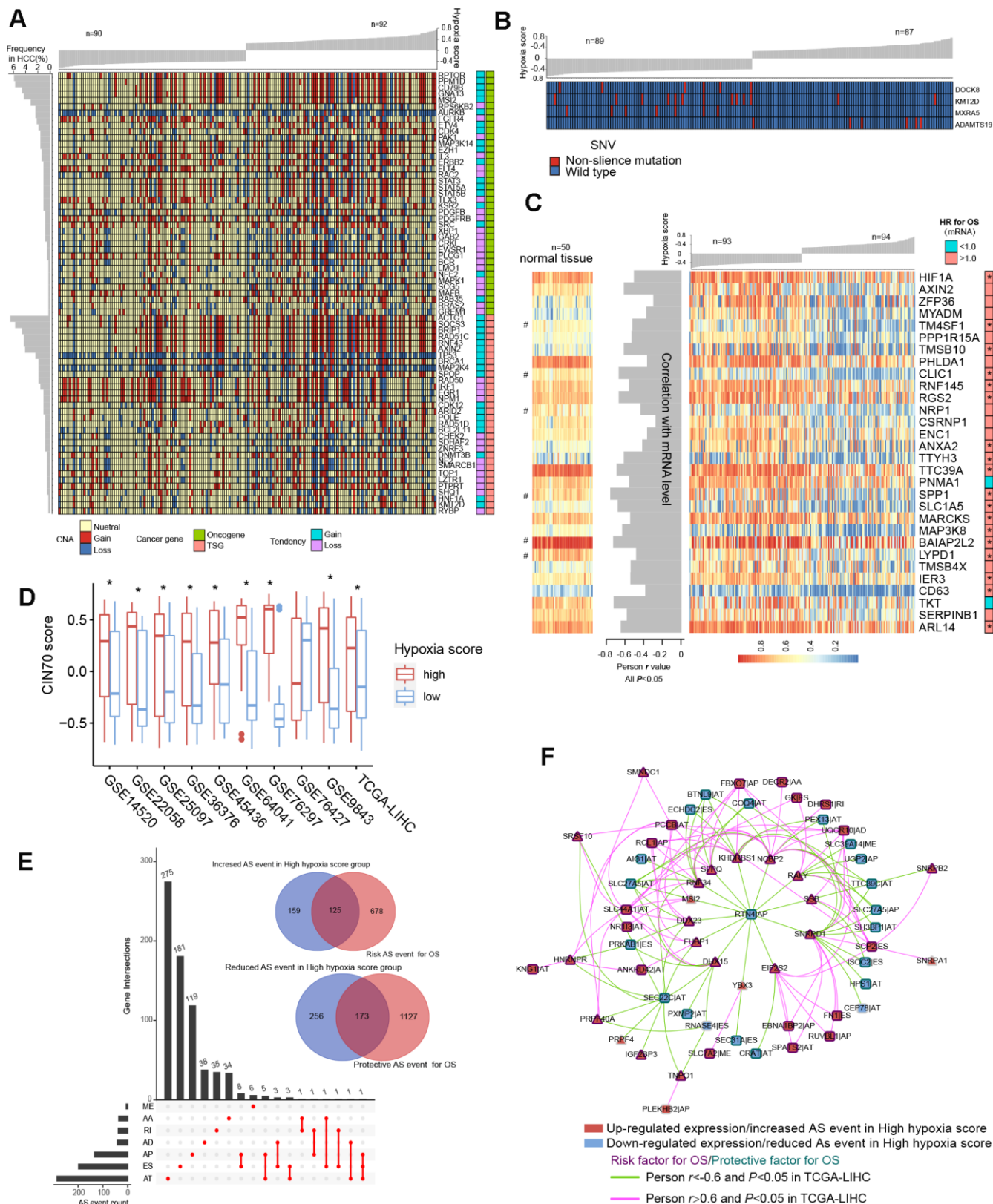
Figure 4. Differences in genomic and epigenetic alterations between groups with high hypoxia scores and low hypoxia scores. (**A**) The difference in the incidence of copy-number aberrations (CNAs) in 71 cancer genes between the high hypoxia score group and the low hypoxia score group. (**B**) The proportions of single-nucleotide variants (SNVs) in 4 genes are significantly different between the high hypoxia score group and the low hypoxia score group. (**C**) Reductions in the methylation levels of 30 genes are accompanied by significant

increases in the corresponding mRNA levels in the high hypoxia score group. The correlation between DNA methylation and corresponding mRNA expression was obtained through Pearson correlation analysis based on TCGA-LIHC data. The hazard ratios (HR) of the corresponding mRNAs for overall survival (OS) were calculated by the logrank test for TCGA-LIHC data, and the cut-off was the median expression level. (**D**) 70-gene chromosome instability (CIN70) was used to assess chromosome instability in tumor tissues from 10 hepatocellular carcinomas (HCC) datasets. The CIN70 scores are significantly different between tumor tissues with high hypoxia scores and low hypoxia scores. (**E**) The occurrences of 713 AS events are significantly different between the high hypoxia score group and the low hypoxia score group. Some of the AS events are associated with the OS of HCC patients. (**F**) The expression of 30 splicing factors in the high hypoxia score group and the low hypoxia score group are different, and their expression trends are consistent in 10 datasets. These splicing factors and the AS events with different occurrences between the two groups form a network. The correlations between the nodes of the network were calculated by Pearson correlation analysis based on TCGA-LIHC data. The relationship between the nodes and the OS of HCC patients was obtained through univariate cox survival analysis.

enrichment analysis of biological processes and pathways. In addition to the HIF1 pathway and sugar and lipid metabolism pathways (indicating their relations to hypoxia exposure), these mRNAs are also involved in various classical pathways of tumorigenesis and development (Supplementary Figure 6A). The PPI network of these mRNA translation products is shown in Supplementary Figure 6B. In addition, we analyzed the differences in expression of 404 splicing factors (SFs) between the high hypoxia score group and the low hypoxia score group. The differences in expression of 30 SFs were relatively consistent in 10 datasets (the condition of adjusted $P < 0.05$ was met in 5 or more datasets). Except for NRIP2, other SFs were upregulated in the high hypoxia score group (Supplementary Figure 7A). Next, correlations among these SFs and 713 AS events were analyzed. When constructing the correlation network, we only set the thresholds at |Person r |> 0.6 and $P < 0.05$ and did not perform other node screening. Notably, in the obtained correlation network, each node reached logical harmony in variation trend and prognostic value (Figure 4F). For example, KHDRBS1 is an SF with elevated expression in the high hypoxia score group and is a risk factor for HCC patient survival. The occurrences of 7 AS events that were negatively correlated with KHDRBS1 were reduced in the high hypoxia score group, and these 7 AS events were protective factors for HCC patient survival. In contrast, 5 AS events that were positively correlated with KHDRBS1 increased in the high hypoxia score group, and these AS events were risk factors for HCC patient survival. Based on the evidence, we believe that AS events may be a response to hypoxia in HCC cells. Some hypoxia-induced changes in cell functions and behaviors may be achieved through the SF-AS network. The difference in AS between the high hypoxia score group and the low hypoxia score group may explain the prognosis difference between the two groups.

RNA N6-methyladenosine (m6A) modification is another important epigenetic feature of RNA [14]. The expression differences in 21 m6A regulators between the high hypoxia score group and the low hypoxia score group were analyzed. In the 10 HCC datasets, most m6A regulators showed no significant difference between the two groups (Supplementary Figure 7B), but 1 m6A writer (WTAP), 1 m6A eraser (ALKBH5), and 1 m6A reader (YTHDF2) significantly increased in the high hypoxia score groups of 4 or more datasets.

Changes in functional proteomics in HCC patients with different hypoxia scores

Based on the reverse phase protein array (RPPA) data of the TCGA-LIHC cohort, we analyzed the differences in the abundance of 218 proteins between the high hypoxia score group and the low hypoxia score group. A total of 50 proteins exhibited differences in abundance between the two groups (Figure 5A). Increases in LKB1 and mTOR activity inhibition are hallmark events of hypoxia exposure, which once again confirmed the efficacy of our 21-gene signature in indicating hypoxia exposure. Approximately half of the protein changes were inconsistent with the changes in their mRNA levels, suggesting that the regulation of posttranslational levels by hypoxia cannot be ignored. The differentially expressed proteins were mainly concentrated in AKT/mTOR and its related pathways. The reduction in the AKT level and its phosphorylation level seems to be a core event. The change trends for some proteins were opposite to those reported by the relevant literature on AKT (Figure 5B), which suggests that in addition to the regulation of AKT, there may be a complex regulatory network that has not been discovered. Survival analysis (cut-off = median value, Figure 5C) showed that a low abundance of AKT, AKT-pT308 and AKT_pS473 indicated worse OS (except for AKT_pS473, logrank $P < 0.05$). Hypoxia may induce some functions by inhibiting AKT in HCC. Overall, the changes in the proteins between the two groups reflected the dynamic balance of autophagy/apoptosis and the response to hypoxic stress to some extent. These proteins may affect the viability of the cells under hypoxia and their tolerance to hypoxia. Moreover, the altered protein abundance between the two groups may be related to the different drug sensitivities of HCC patients in the two groups. The

correlation between some proteins differentially expressed in the two groups and drug sensitivity is shown in Supplementary Figure 8 (based on data from the MD Anderson Cell Lines Project and the Genomics of Drug Sensitivity in Cancer database). For example, the reduction in AKT_pS473 may cause the sensitivity of cells to sorafenib to decrease. Therefore, when formulating a treatment plan, it may be necessary to evaluate the hypoxic condition of the lesions.

DISCUSSION

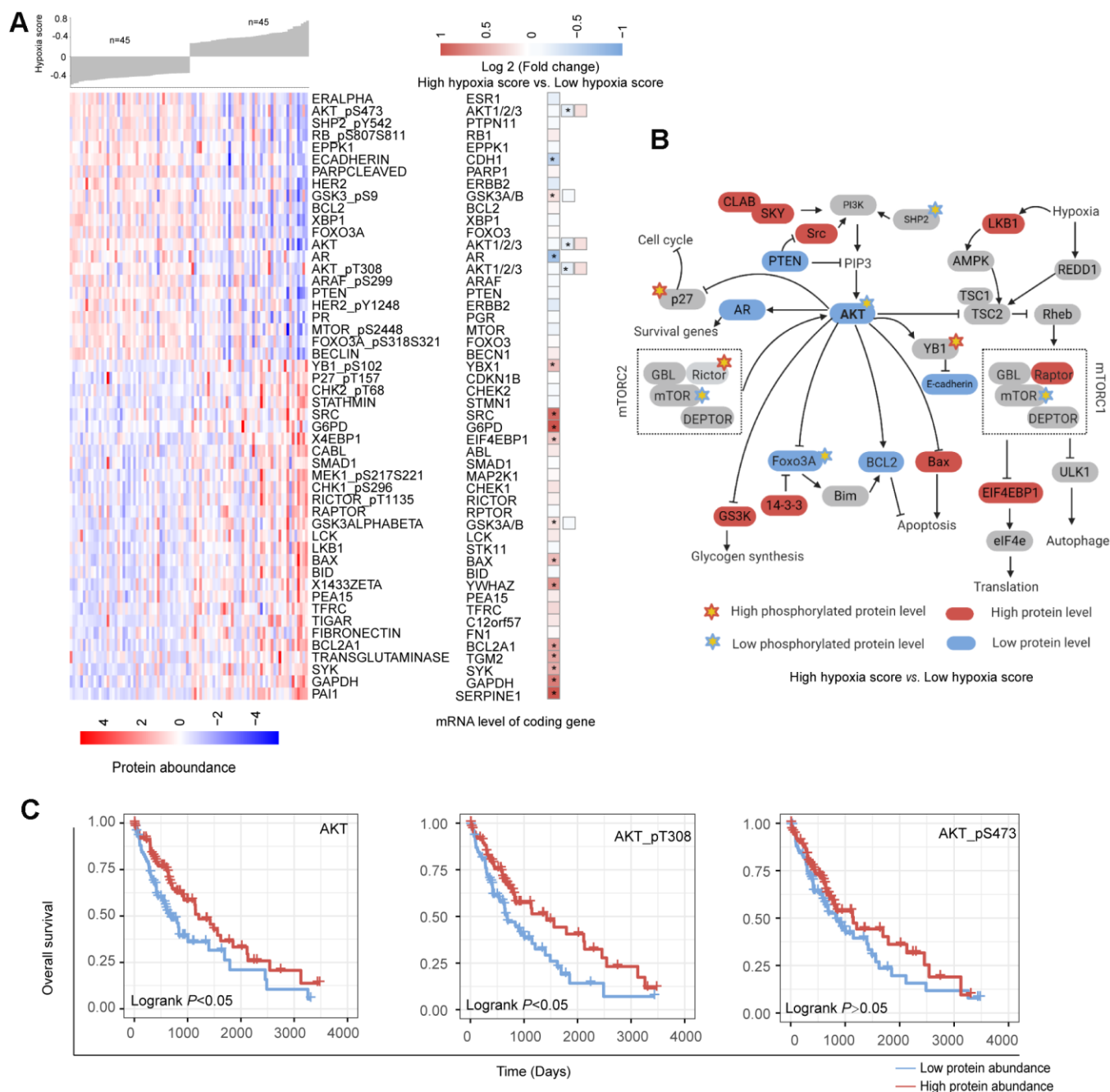The molecular changes in cancer cells caused by hypoxia have been widely reported in studies *in vitro* [15, 16].

Figure 5. Changes in functional proteomics between hepatocellular carcinoma (HCC) patients with high hypoxia scores and low hypoxia scores. (**A**) Fifty proteins with significant differences in abundance between HCC patients with high hypoxia scores and low hypoxia scores. (**B**) A schematic diagram of some proteins with significant differences in abundance in the AKT/mTOR pathway. (**C**) The low abundance of AKT, AKT-pT308, and AKT_pS473 proteins in TCGA-LIHC indicates poor overall survival.

However, the molecular changes caused by hypoxia in tissues may be different from the data obtained *in vitro* [17]. If it is necessary for patients to truly benefit from anti-hypoxic treatment strategies, only the data obtained from the cellular hypoxia model is far from enough. Polarographic electrodes can accurately measure the degree of hypoxia exposure in cancer tissues, but this method does not have enough practicality [18]. Hypoxia can change the gene expression level in tumor cells. Therefore, the detection of the levels of specific genes can indirectly reflect hypoxia exposure in tissues [19]. In the past decade, different hypoxia gene signatures have been reported, and some excellent signatures, such as Buffa's 15-gene hypoxia signature, have proven to have the ability to indicate hypoxia in a variety of tumors [20]. Sorensen reported a group of hypoxia gene markers indicating the hypoxia in human squamous cell carcinomas [21]. Inspired by these studies we have identified a novel 21-gene hypoxia signature that has excellent robustness in the assessment of hypoxia exposure. This provides a basis for us to explore the hypoxia-induced molecular landscape in HCC. After calculating hypoxia scores using the 21-gene hypoxia signature, we found that the hypoxia scores in the tumor tissues of HCC patients in 11 datasets were grouped into 2 clusters with different hypoxia exposure levels. Because the number of probes in the GSE10141 dataset was small, this dataset was excluded in subsequent studies. Considering the numerical distribution of hypoxia scores in tissues, we selected the upper quartile and the lower quartile to divide high hypoxia scores and low hypoxia scores. In this way, more differential changes may be obtained, creating a situation that is more in line with clinical practice. Because hypoxic conditions and molecular changes in tumor tissues are nonlinearly correlated, mild hypoxia exposure may not cause many molecular events. After comprehensive analysis in the present study, the patients in the high hypoxia score group and the patients in the low hypoxia score group were found to have many transcriptomic, genomic, epigenomic, and proteomic differences.

Among HF/DE-mRNAs, all risk factors for OS were significantly upregulated in the high hypoxia score group while the protective factors were downregulated; this phenomenon occurred in almost all cohorts. For example, ANXA5 and SPP1 were both risk factors for OS in the TCGA-LIHC and GSE14520 datasets and were significantly overexpressed in the high hypoxia score group in 9 cohorts. The results of pathway enrichment analysis indicated the P53 pathway, PI3K/AKT, (mTOR) signaling, and TGF-beta signaling upregulated in the high hypoxia group. The activation of these pathways is critical to the onset of the progression of tumors and can be used to explain the

poor prognosis of patients in the high hypoxia score group. In addition, many drugs target these pathways, suggesting that the degree of hypoxia exposure needs to be considered in the selection of medications for patients with HCC. Among DE-miRs, miR-210 is a frequently reported master hypoxia-responsive miR, and its high expression has been observed in a variety of hypoxia-treated tumor cells [22, 23]. The presence of miR-210 in the DE-miR list means that our 21-gene hypoxia signature can reflect hypoxia exposure in HCC. We also obtained DE-miRs-HF/DE-mRNAs networks. These networks suggested that the changes in some DE-mRNAs in HCC tissue under hypoxic conditions might be caused by the targeted regulation of DE-miRs. For example, miR-216b-5p is the miR with the highest degree of reduced expression in the high hypoxia score group (log2FC = -2.61, adjusted P < 0.001), and it is complementary to the 3' untranslated region (3'-UTR) of HK2. The prediction was supported by argonaute-crosslinking immunoprecipitation sequencing (AGO-CLIP-seq) data [24, 25]. HK2 is a definite hypoxia-inducible gene [26]. The above results suggest that the high expression of HK2 in hypoxia may be associated with the reduction in miR-216b-5p. These 2 networks provide additional explanations for hypoxia-induced transcriptome changes in HCC tissues in addition to the HIF-1A-related mechanism. Attention needs to be paid to miRs with high connectivity in the network, such as miR-1224-5p, miR-877-5p, let-7a-2-3p, and miR-378c. DE-lncRNA was also analyzed as well as mRNAs and miRs. Moreover, we constructed a DE-lncRNA-DE-microRNA-HF/DE-mRNA network. There were differences in the expression of mRNAs, miRs, and lncRNAs between the high hypoxia score and low hypoxia score groups. Many cancer-promoting RNA molecules were upregulated in the high hypoxia score group while cancer-suppressing RNA molecules were downregulated in the high hypoxia score group. Some DE-mRNAs and/or DE-lncRNAs and DE-mRNAs may form regulatory networks that participate in the development of HCC and affect the prognosis of HCC patients.

In genomic alterations analysis, CNAs changes suggested that the genomic instability of HCC patients may be related to hypoxia exposure. Few genes were significantly different in the incidence of SNVs between the high hypoxia and low hypoxia score groups. Hence, we believed that hypoxia has a limited ability to induce SNVs, and important SNVs likely occur before the onset of hypoxia. Epigenetic alterations suggested that hypoxia exposure may promote the hypomethylation of some genes related to the onset and progression of HCC and that the hypomethylation of some other genes may be a characteristic feature of hypoxia exposure. Moreover, RNA alternative splicing and m6A

modification were another important epigenetic feature under hypoxia which expanded our understanding of the role of hypoxia in HCC.

Some differential molecular events identified in the present study are closely related to the prognosis of patients and are molecular mechanisms that explain the cancer-promoting effect of hypoxia. Therefore, these differential events can help screen HCC diagnostic and treatment targets. The results at multi-omic levels had some consistency, such as the simultaneous presence of high mRNA expression and decreased methylation of the corresponding DNA in the high hypoxia score group. However, some results were contradictory. For example, the GSEA results for the AKT/mTOR pathway based on the mRNA level were not consistent with the proteomic results for the AKT/mTOR pathway. Therefore, it is necessary to investigate the molecular changes caused by hypoxia using multi-omics approaches. Based on the available data, we analyzed the relationships between differential events and patient prognosis and extracted useful data for translational medicine. In the Supplementary Files, we provided reference data with details, such as the lists of DE-mRNAs, DE-miRNAs, and DE-lncRNAs, the list of genes with CNAs, the list of genes with aberrant methylation loci, and the list of hypoxia-related RNA alternative splicing events. We hope that these data can inspire and help other researchers, improve research efficiency, and narrow the scope of research.

It should be emphasized that we found that the HIF-1A mRNA levels significantly increased in the high hypoxia score groups in the 10 datasets and showed a significantly positive correlation with the hypoxia score. The reason why this point is emphasized is that most previous studies have focused on the posttranslational regulation of HIF-1A at the protein level, and some researchers even deliberately ignore changes in HIF-1A mRNA [27]. Our data suggest that the regulation of HIF-1A mRNA levels cannot be ignored in hypoxia exposure, meaning that in treatment regimens targeting HIF-1A, not only HIF-1A protein but also HIF-1A mRNA should be targeted. In addition, our data showed that during hypoxia exposure, the level of HIF-1A might be regulated by a multilevel positive feedback network and that an mRNA/lncRNA network and methylation regulation might be part of this network.

In summary, the present study provided tissue-level evidence for throwing light on the mechanism by which hypoxia may perform its function in HCC. An unprecedented comprehensive panorama of hypoxia associated multi-omics molecular features was systematically described. Networks between molecular features were tried to establish such as DE-lncRNA-

DE-microRNA-HF/DE-mRNA networks. Clinically, the data obtained from this study may facilitate developing possible treatment or diagnostic novel targets under the perspective of personalized medicine.

MATERIALS AND METHODS

Public datasets in GEO

We retrieved three independent mRNA microarray datasets (GSE18494, GSE55214, and GSE57613) based on hypoxia treated HCC cells from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). Datasets and available clinical information of ten HCC cohorts (GSE14520, GSE22058, GSE25097, GSE36376, GSE45436, GSE64041, GSE76297, GSE76427, GSE10141 and GSE9843) were downloaded from GEO. All gene symbols in GEO datasets were converted to the latest HGNC (HUGO Gene Nomenclature Committee) symbols.

Multi-omic data and clinical data of TCGA-LIHC

mRNA expression, microRNA expression, lncRNA expression, CNAs, SNVs, DNA methylation data of the Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) were downloaded from the TCGA data portal (https://portal.gdc.cancer.gov/). The RNA alternative splicing data of TCGA-LIHC were obtained from TCGA SpliceSeq (http://projects.insilico.us.com/TCGASpliceSeq/). The reverse phase protein array (RPPA) data of TCGA-LIHC were downloaded from the cancer proteome atlas (https://www.tcpaportal.org/tcpa/index.html).

Hypoxia scores calculation

In our previous study, a novel 21-gene hypoxia signature was built using robust rank aggregation (RRA) algorithm from the microarray data of hypoxia treated HUH7, SNU-182, and HLF cells. Hypoxia score was calculated using gene set variation analysis (GSVA) based on the 21 genes. The gene lists of 21-gene signature were shown in Additional File 1: Supplementary Table 1.

Biological process and pathway enrichment assay

The biological process and pathway enrichment assay of candidate genes were performed using online tools provided by Metascape (http://metascape.org/gp/index.html). The enrichment analysis has been carried out with the following ontology sources: the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway, Gene Ontology (GO) biological processes, Reactome Gene Sets, and Canonical Pathways All

genes in the genome have been used as the enrichment background. To establish the network based on the relationships between the terms, a subset of enriched terms with a similarity > 0.3 were connected by edges. 20 clusters were obtained and the terms with the best p-values were selected (for more details see the website of Metascape). The network was visualized using Cytoscape (version 3.7.2).

Protein-protein interaction enrichment analysis

We used online tools provided by Metascape to perform protein-protein interaction enrichment analysis for the production of candidate genes. According to data from BioGrid, InWeb, and OmniPath, a resultant network contains the subset of proteins that form physical interactions with at least one other member was built using Molecular Complex Detection (MCODE) algorithm [28]. Then pathway and process enrichment analysis has been applied to each MCODE component (for more details see the website of Metascape). The network was visualized using Cytoscape (version 3.7.2).

Gene set enrichment analysis (GSEA)

GSEA was performed for candidate mRNAs across HCC cohorts using GSEA tools (version 4.0.3) provided by the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/index.jsp). The hallmark gene was set as the reference genes. The significantly activated or suppressed pathways were identified as pathways with P value<0.05 and FDR<0.25.

Interactions analysis for mRNA/microRNA/lncRNA

miRNA-target interactions were presented by intersecting the predicting target sites of miRNAs with binding sites of Ago protein using the Encyclopedia of RNA Interactomes (ENCORI, http://starbase.sysu.edu.cn/index.php) and miRwalk 3.0 (http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk/). The miRNA-lncRNA interactions were presented using LncBase v.2 experimental module tools (http://diana.imis.athena-innovation.gr/DianaTools/index.php). Interactions of mRNA/microRNA/lncRNA were visualized using Cytoscape (version 3.7.2).

Statistical analysis

Statistical analyzes were performed using R software (version 3.6.1) with relevant packages. In brief, the differential expressed mRNAs were extracted from microarray datasets using the Limma package. The differential expressed mRNAs, microRNA, and lncRNA of TCGA-LIHC were identified using Linnorm packages.

The difference between the two groups was compared using the independent t-test or Wilcox test. The adjusted P value was obtained using the false discovery rate (FDR) method. Coefficients were calculated using Pearson or Spearman's correlation analysis. A chi-squared test was used to determine the significant difference between the frequencies. Survival analysis was performed using Univariate Cox/multivariate analysis hazard analysis or Kaplan-Meier survival estimate using the survival package. The forest-plot R package was employed to visualize the hazard rate obtained from survival analysis. The Kaplan-Meier survival curves were created using the survminer package with the logrank test. In the present study, statistical significance was set at a probability value of $P < 0.05$.

REFERENCES

1. Xiong XX, Qiu XY, Hu DX, Chen XQ. Advances in hypoxia-mediated mechanisms in hepatocellular carcinoma. Mol Pharmacol. 2017; 92:246–55.
   https://doi.org/10.1124/mol.116.107706
   PMID:28242743

2. Wong CC, Kai AK, Ng IO. The impact of hypoxia in hepatocellular carcinoma metastasis. Front Med. 2014; 8:33–41.
   https://doi.org/10.1007/s11684-013-0301-3
   PMID:24234682

3. Méndez-Blanco C, Fondevila F, García-Palomo A, González-Gallego J, Mauriz JL. Sorafenib resistance in hepatocarcinoma: role of hypoxia-inducible factors. Exp Mol Med. 2018; 50:1–9.

https://doi.org/10.1038/s12276-018-0159-1
PMID:30315182

4. Lencioni R. Chemoembolization for hepatocellular carcinoma. Semin Oncol. 2012; 39:503–09.
https://doi.org/10.1053/j.seminoncol.2012.05.004
PMID:22846867

5. Petrillo M, Patella F, Pesapane F, Suter MB, Ierardi AM, Angileri SA, Floridi C, de Filippo M, Carrafiello G. Hypoxia and tumor angiogenesis in the era of hepatocellular carcinoma transarterial loco-regional treatments. Future Oncol. 2018; 14:2957–67.
https://doi.org/10.2217/fon-2017-0739
PMID:29712486

6. Severi T, van Malenstein H, Verslype C, van Pelt JF. Tumor initiation and progression in hepatocellular carcinoma: risk factors, classification, and therapeutic targets. Acta Pharmacol Sin. 2010; 31:1409–20.
https://doi.org/10.1038/aps.2010.142 PMID:20953207

7. van Malenstein H, Gevaert O, Libbrecht L, Daemen A, Allemeersch J, Nevens F, Van Cutsem E, Cassiman D, De Moor B, Verslype C, van Pelt J. A seven-gene set associated with chronic hypoxia of prognostic importance in hepatocellular carcinoma. Clin Cancer Res. 2010; 16:4278–88.
https://doi.org/10.1158/1078-0432.CCR-09-3274
PMID:20592013

8. Bhandari V, Hoey C, Liu LY, Lalonde E, Ray J, Livingstone J, Lesurf R, Shiah YJ, Vujcic T, Huang X, Espiritu SM, Heisler LE, Yousif F, et al. Molecular landmarks of tumor hypoxia across cancer types. Nat Genet. 2019; 51:308–18.
https://doi.org/10.1038/s41588-018-0318-2
PMID:30643250

9. Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, Ruan H, Zhang Z, Song A, Zhang H, Liu L, Diao L, Lou Y, et al. Characterization of hypoxia-associated molecular features to aid hypoxia-targeted therapy. Nat Metab. 2019; 1:431–44.
https://doi.org/10.1038/s42255-019-0045-8
PMID:31984309

10. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, Leiserson MD, Niu B, McLellan MD, Uzunangelov V, Zhang J, Kandoth C, Akbani R, et al, and Cancer Genome Atlas Research Network. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell. 2014; 158:929–44.
https://doi.org/10.1016/j.cell.2014.06.049
PMID:25109877

11. Eferl R, Trauner M. Chromosomal instability in HCC: a key function for checkpoint kinase 2. Gut. 2018; 67:204–05.

https://doi.org/10.1136/gutjnl-2017-314101
PMID:28476916

12. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet. 2006; 38:1043–48.
https://doi.org/10.1038/ng1861 PMID:16921376

13. Lu Y, Xu W, Ji J, Feng D, Sourbier C, Yang Y, Qu J, Zeng Z, Wang C, Chang X, Chen Y, Mishra A, Xu M, et al. Alternative splicing of the cell fate determinant numb in hepatocellular carcinoma. Hepatology. 2015; 62:1122–31.
https://doi.org/10.1002/hep.27923
PMID:26058814

14. Chen Y, Peng C, Chen J, Chen D, Yang B, He B, Hu W, Zhang Y, Liu H, Dai L, Xie H, Zhou L, Wu J, Zheng S. WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. Mol Cancer. 2019; 18:127.
https://doi.org/10.1186/s12943-019-1053-8
PMID:31438961

15. Nam HY, Chandrashekar DS, Kundu A, Shelar S, Kho EY, Sonpavde G, Naik G, Ghatalia P, Livi CB, Varambally S, Sudarshan S. Integrative epigenetic and gene expression analysis of renal tumor progression to metastasis. Mol Cancer Res. 2019; 17:84–96.
https://doi.org/10.1158/1541-7786.MCR-17-0636
PMID:30131446

16. Dong ZZ, Yao M, Wang L, Wu W, Gu X, Yao DF. Hypoxia-inducible factor-1alpha: molecular-targeted therapy for hepatocellular carcinoma. Mini Rev Med Chem. 2013; 13:1295–304.
https://doi.org/10.2174/1389557511313090004
PMID:22512580

17. Saxena K, Jolly MK. Acute vs. Chronic vs. Cyclic hypoxia: their differential dynamics, molecular mechanisms, and effects on tumor progression. Biomolecules. 2019; 9:339.
https://doi.org/10.3390/biom9080339
PMID:31382593

18. Rickard AG, Palmer GM, Dewhirst MW. Clinical and pre-clinical methods for quantifying tumor hypoxia. Adv Exp Med Biol. 2019; 1136:19–41.
https://doi.org/10.1007/978-3-030-12734-3_2
PMID:31201714

19. Eustace A, Mani N, Span PN, Irlam JJ, Taylor J, Betts GN, Denley H, Miller CJ, Homer JJ, Rojas AM, Hoskin PJ, Buffa FM, Harris AL, et al. A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. Clin Cancer Res. 2013; 19:4879–88.

https://doi.org/10.1158/1078-0432.CCR-13-0542
PMID:23820108

20. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. Br J Cancer. 2010; 102:428–35.
https://doi.org/10.1038/sj.bjc.6605450
PMID:20087356

21. Sørensen BS, Toustrup K, Horsman MR, Overgaard J, Alsner J. Identifying pH independent hypoxia induced genes in human squamous cell carcinomas *in vitro*. Acta Oncol. 2010; 49:895–905.
https://doi.org/10.3109/02841861003614343
PMID:20429727

22. Dang K, Myers KA. The role of hypoxia-induced miR-210 in cancer progression. Int J Mol Sci. 2015; 16:6353–72.
https://doi.org/10.3390/ijms16036353
PMID:25809609

23. Huang X, Le QT, Giaccia AJ. MiR-210—micromanager of the hypoxia pathway. Trends Mol Med. 2010; 16:230–37.
https://doi.org/10.1016/j.molmed.2010.03.004
PMID:20434954

24. Kishore S, Jaskiewicz L, Burger L, Hausser J, Khorshid M, Zavolan M. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods. 2011; 8:559–64.

https://doi.org/10.1038/nmeth.1608
PMID:21572407

25. Xue Y, Ouyang K, Huang J, Zhou Y, Ouyang H, Li H, Wang G, Wu Q, Wei C, Bi Y, Jiang L, Cai Z, Sun H, et al. Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell. 2013; 152:82–96.
https://doi.org/10.1016/j.cell.2012.11.045
PMID:23313552

26. Xu F, Yan JJ, Gan Y, Chang Y, Wang HL, He XX, Zhao Q. miR-885-5p negatively regulates warburg effect by silencing hexokinase 2 in liver cancer. Mol Ther Nucleic Acids. 2019; 18:308–19.
https://doi.org/10.1016/j.omtn.2019.09.002
PMID:31614321

27. Lin D, Wu J. Hypoxia inducible factor in hepatocellular carcinoma: a therapeutic target. World J Gastroenterol. 2015; 21:12171–78.
https://doi.org/10.3748/wjg.v21.i42.12171
PMID:26576101

28. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003; 4:2.
https://doi.org/10.1186/1471-2105-4-2
PMID:12525261
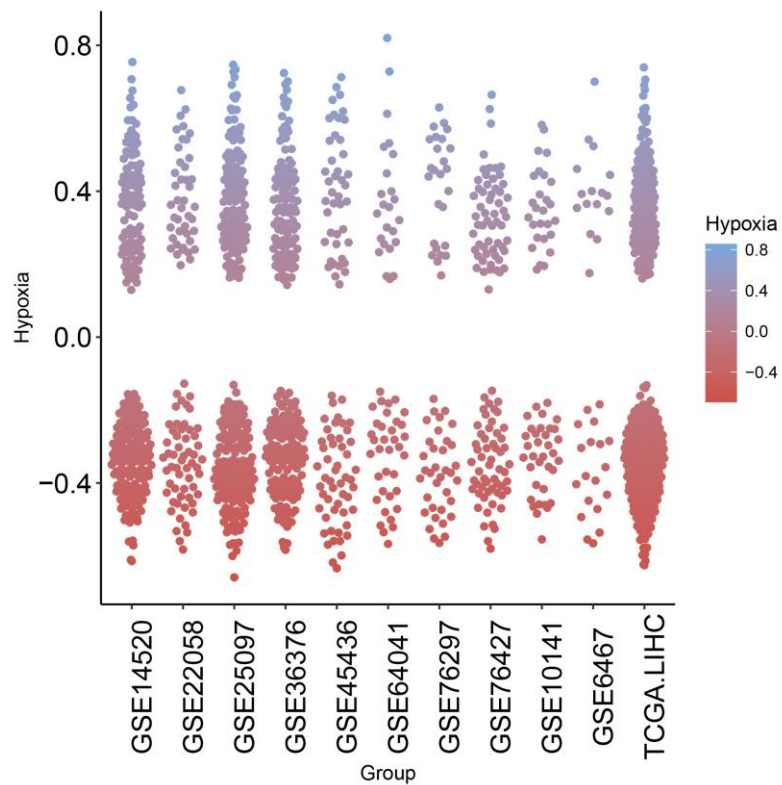
Supplementary Figure 1. The distribution of hypoxia scores calculated based on the 21-gene hypoxia signature in the cancer tissues of 11 HCC cohorts.

upregulated DE-miRs/downregulated DE-mRNAs
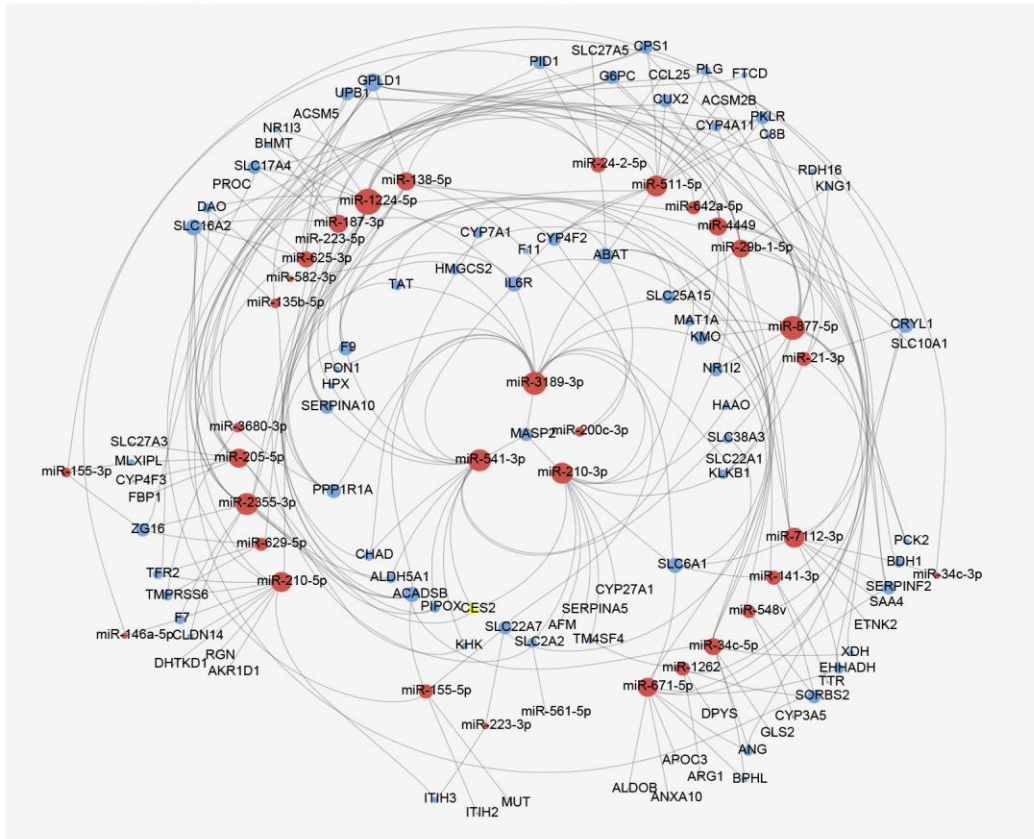
downregulated DE-miRs/upregulated DE-mRNAs
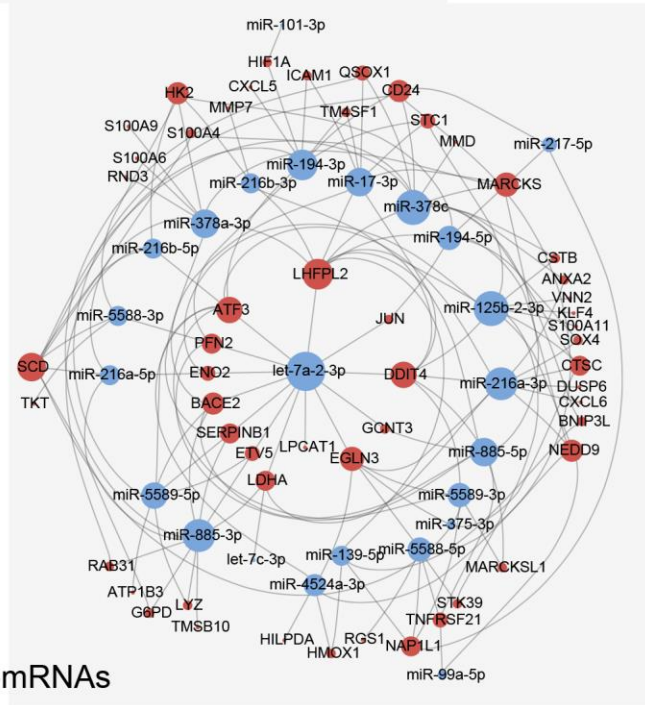
Supplementary Figure 2. DE-miRNAs and HF/DE-mRNAs form a target interaction network. Target interactions were determined by intersecting the predicted target sites of miRNAs with binding sites of Argonaute protein using the online tools ENCORI and miRwalk 3.0.
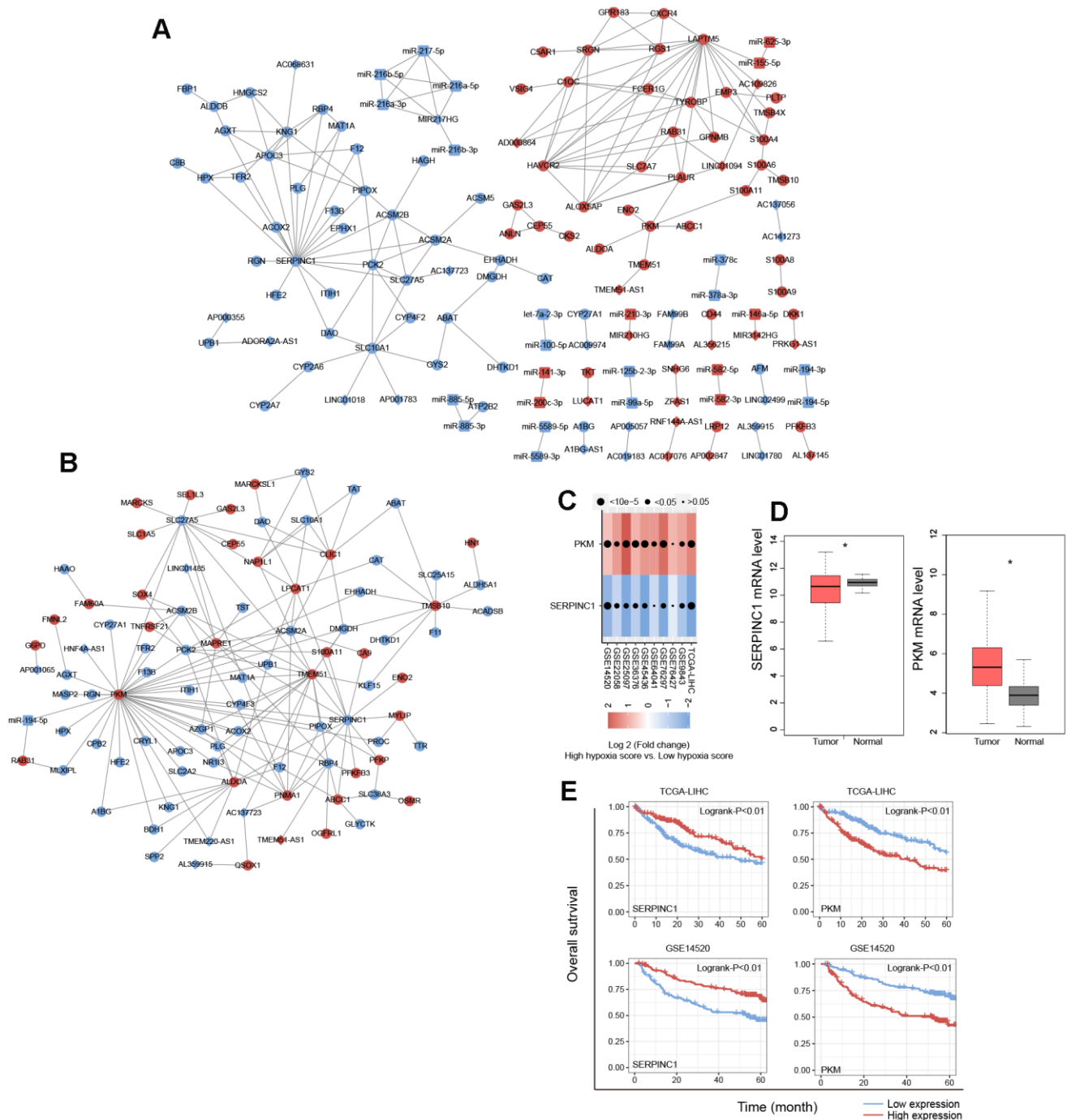
Supplementary Figure 3. A co-expression network formed by HF/DE-mRNAs, DE-miRs, and DE-lncRNAs. (**A**, **B**) The correlations among HF/DE-mRNAs, DE-miRs, and DE-lncRNAs were obtained using Pearson correlation analysis. The nodes with |Pearson r| > 0.8 and P < 0.05 constitute positive and negative co-expression networks. (**C**) In the positive and negative co-expression networks, the 2 genes with the highest connectivity are SERPINC1 and PKM. Their differences among 10 hepatocellular carcinomas (HCC) datasets are shown in the heat map. (**D**) Differences in the expression of SERPINC1 and PKM between normal tissues and cancer tissues of HCC patients in TCGA-LIHC. (**E**) SERPINC1 and PKM are associated with patient survival in TCGA-LIHC and GSE14520.
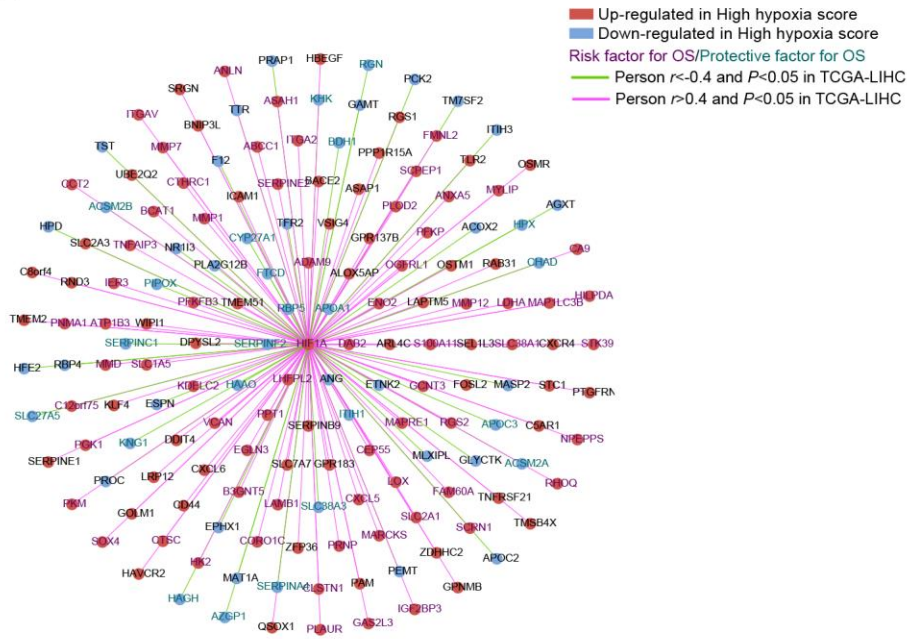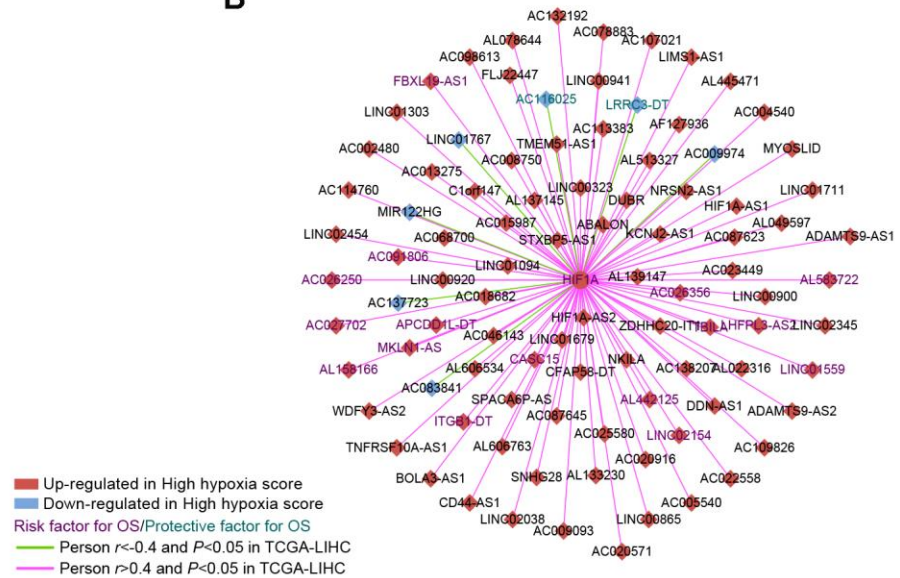
Supplementary Figure 4. DE-mRNAs and DE-lncRNAs correlated with hypoxia-inducible factor 1-alpha (HIF-1A) mRNA. The correlations were calculated by Pearson correlation analysis using TCGA-LIHC data. (**A**) DE-mRNAs correlated with HIF-1A mRNA. (**B**) DE-lncRNAs correlated with HIF-1A mRNA.
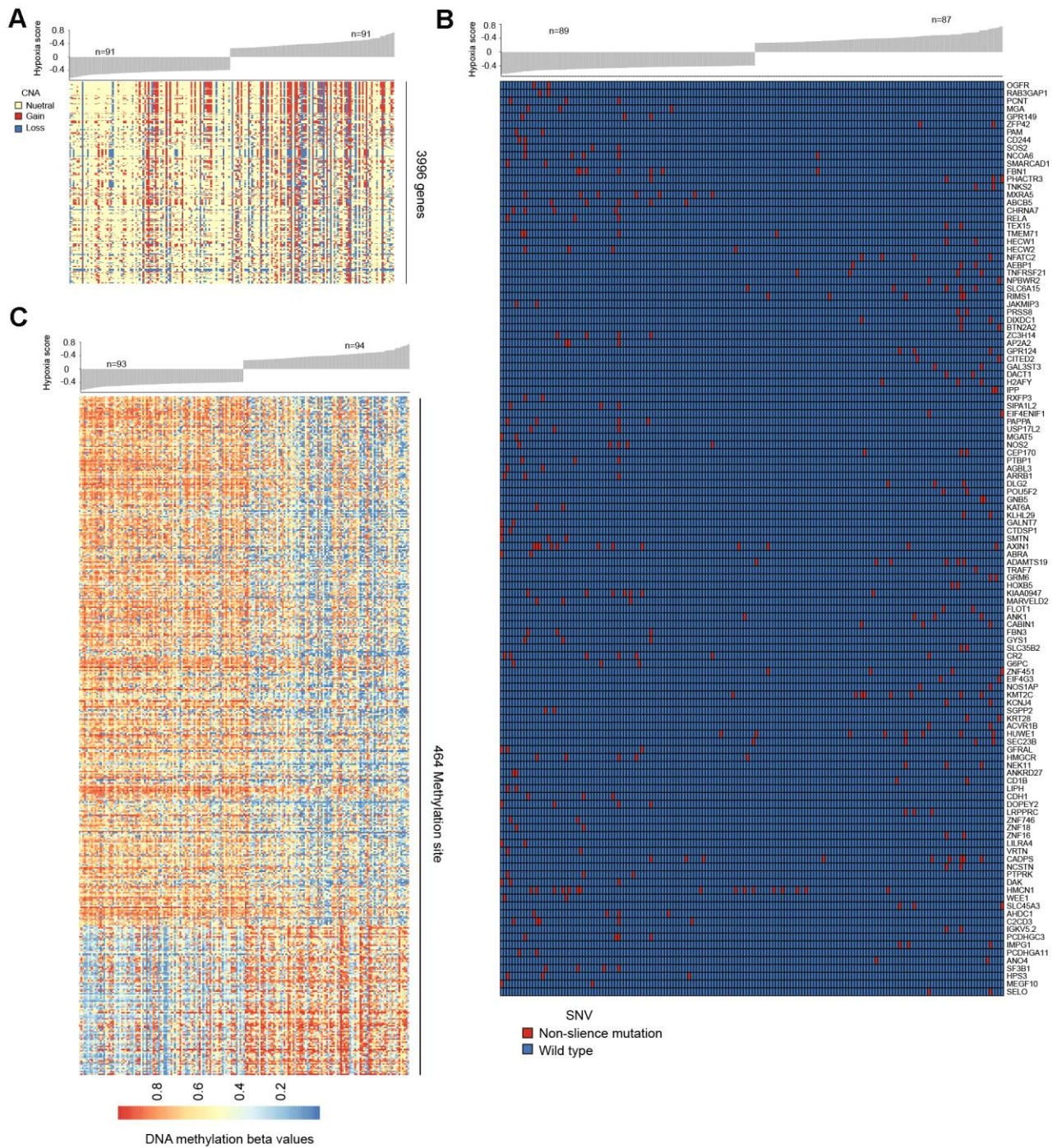
Supplementary Figure 5. Overview of genomic alterations and DNA methylation changes in groups with high hypoxia scores and low hypoxia scores. (**A**) Difference in the proportions of copy-number aberrations (CNAs) between the high hypoxia score group and the low hypoxia score group. (**B**) Proportions of single-nucleotide variants (SNVs) in 172 genes distributed in the low hypoxia score group and the high hypoxia score group. (**C**) A total of 464 loci with different methylation levels in the high hypoxia score group and the low hypoxia score group.
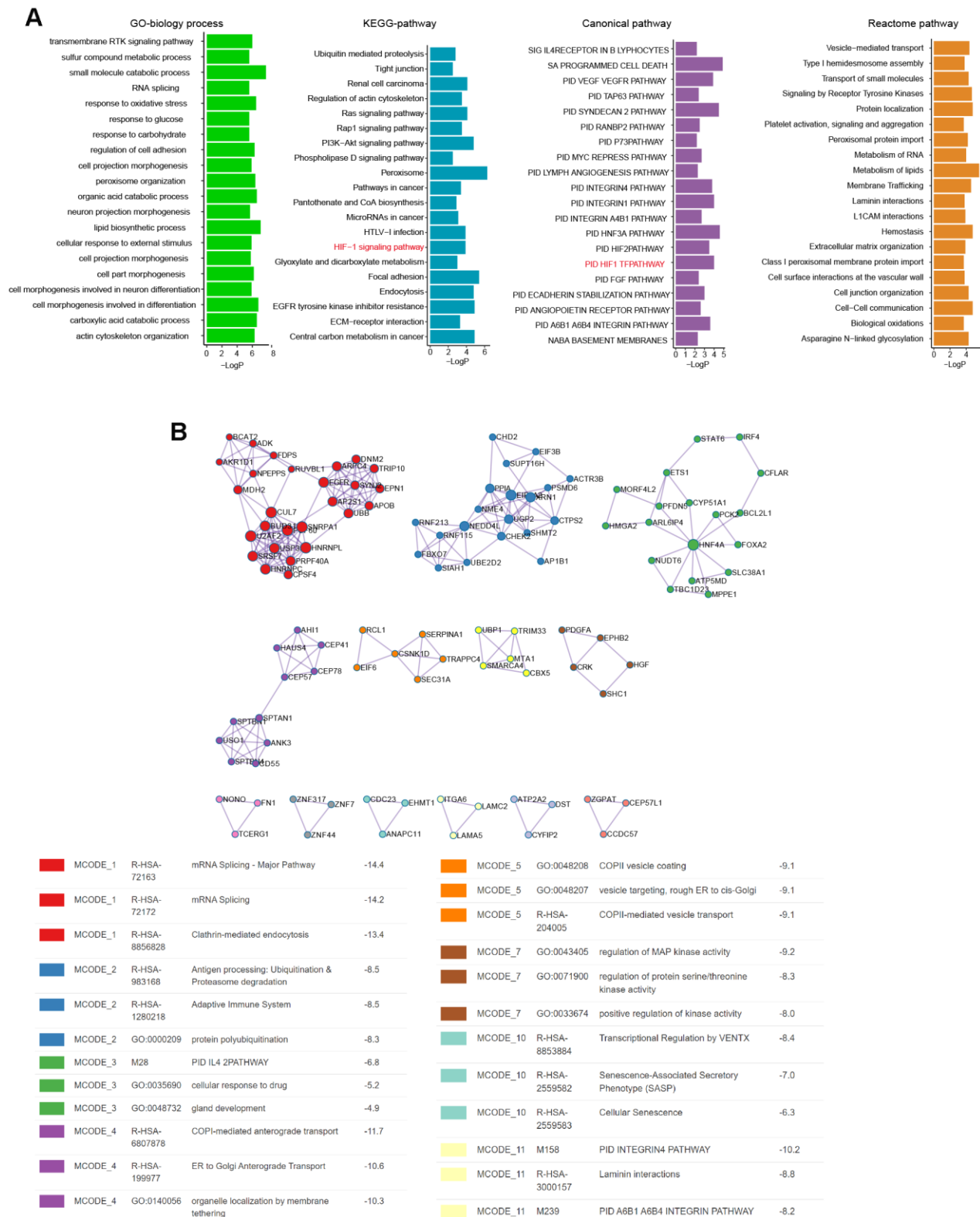
Supplementary Figure 6. Functional enrichment analysis of biological processes and pathways for 681 mRNAs involving differential AS events. (**A**) A total of 681 mRNAs are involved in AS events with differences in occurrences between the high hypoxia score group and the low hypoxia score group. The top 20 (sorted by P-value) enrichment analysis results for GO biological processes and enrichment analysis results for pathways from 3 different data sources are displayed. (**B**) The protein-protein interaction (PPI) enrichment network of the translation products of the 681 mRNAs were constructed using the MCODE algorithm.
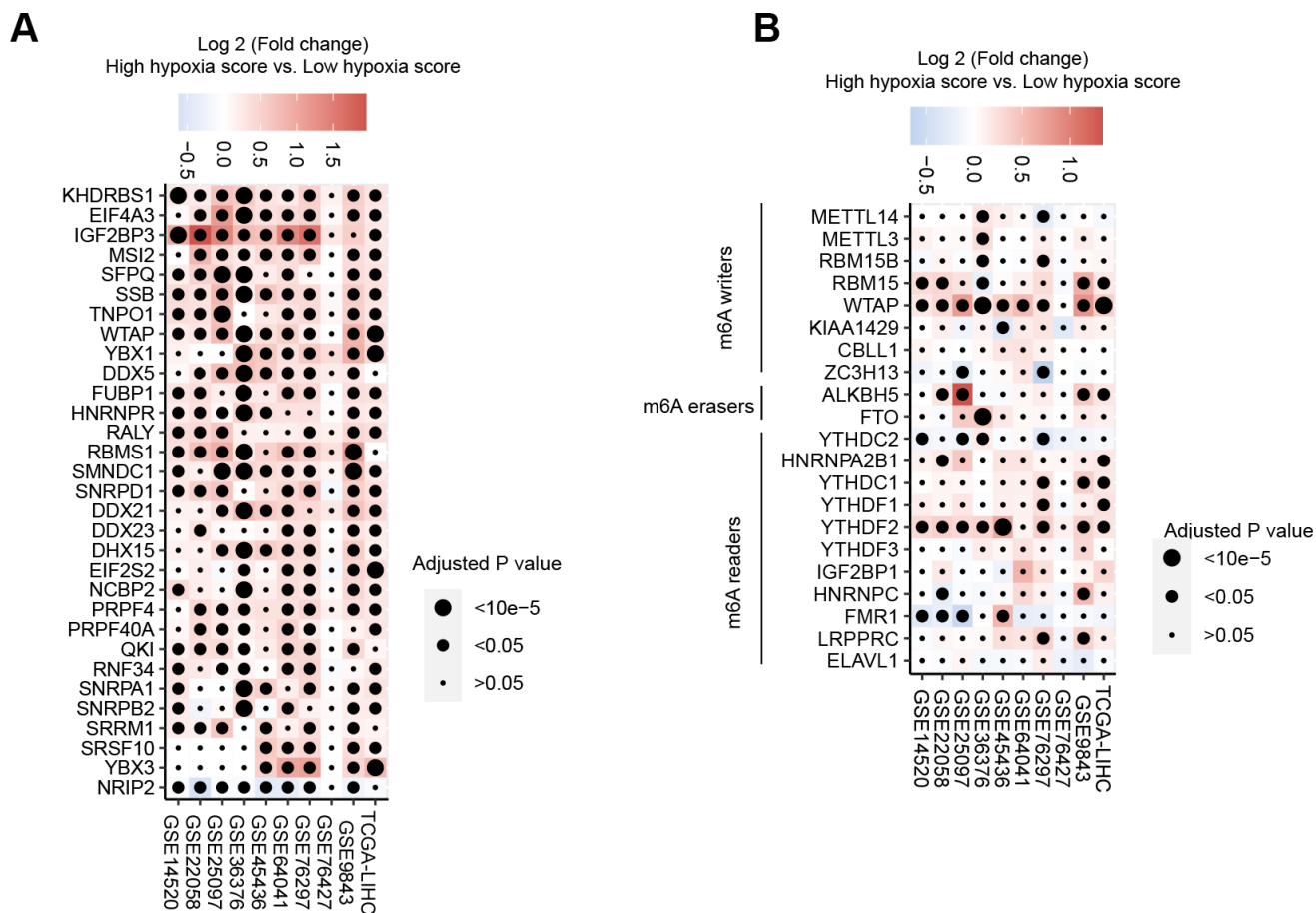
Supplementary Figure 7. The expression of 30 splicing factors and 21 N6-methyladenosine (m6A) regulators in the high hypoxia score group and the low hypoxia score group. (**A**) Differences in the expression of 30 splicing factors between the high hypoxia score groups and the low hypoxia score groups in 10 HCC datasets. (**B**) Differences in the expression of 21 m6A regulators between the high hypoxia score groups and the low hypoxia score groups in 10 HCC datasets.
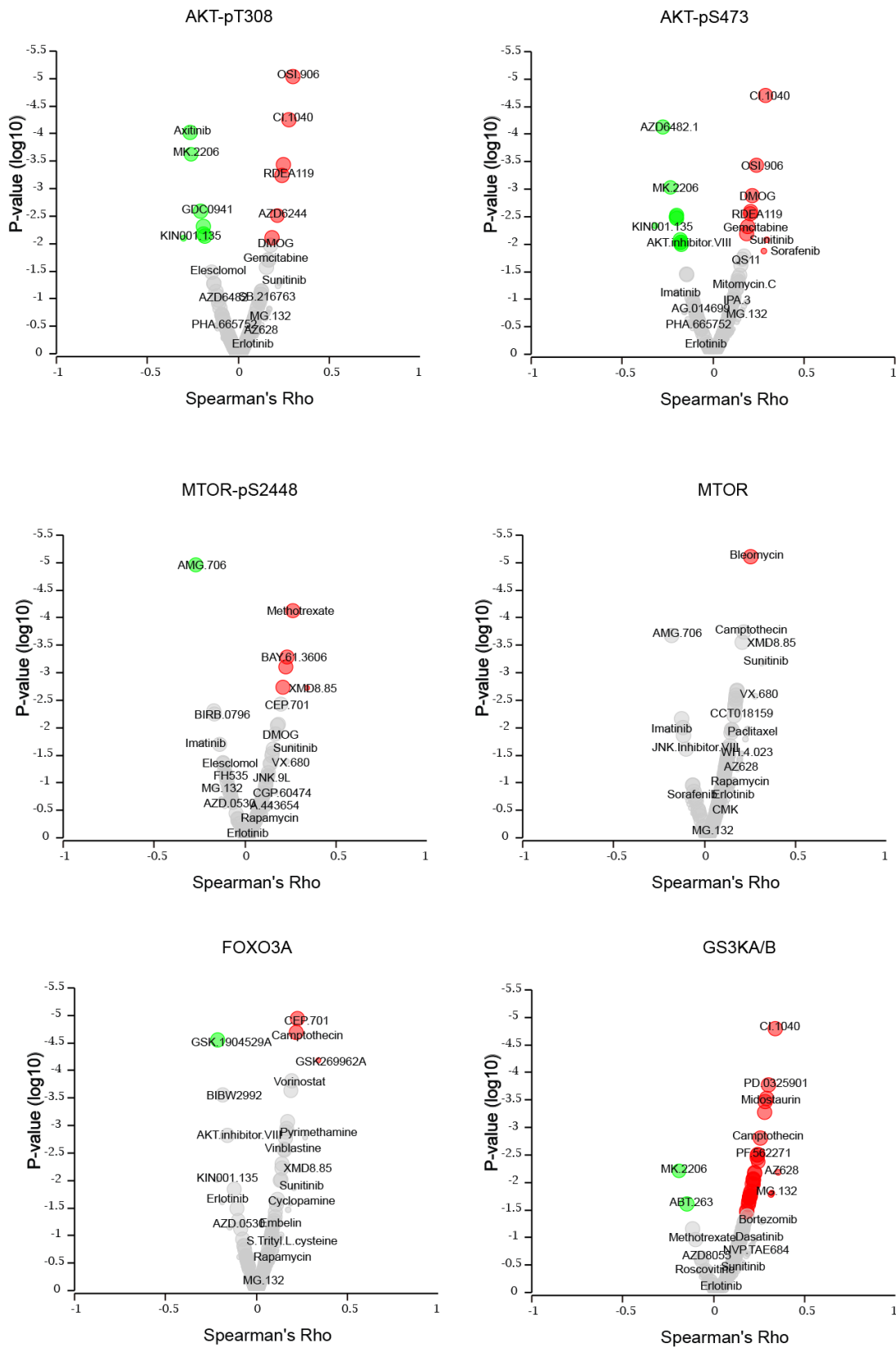
Supplementary Figure 8. Spearman's correlation coefficients for proteins with significantly different expression in the high hypoxia score group and the low hypoxia score group and commonly used anticancer drugs. The data were obtained from the MD Anderson Cell Lines Project and the Genomics of Drug Sensitivity in Cancer database.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 2–4, 7–10, 12.

Supplementary Table 1. The information for elements of 21-gene hypoxia signature.

Gene symbol	Description	Gene Location	Protein Function (Protein Atlas)	Refseq
CA9	carbonic anhydrase 9	chr9:35673928-35681157	Enzymes/ENZYME proteins/Lyases	NM_001216
PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	chr3:48517684-48556803;chr3:48517684-48561129	Enzymes/ENZYME proteins/{Hydrolases,Transferases}	NM_001317134
HILPDA	hypoxia inducible lipid droplet associated	chr7:128455878-128458418;chr7:128455830-128458418	N/A	NM_001098786
BNIP3L	BCL2 interacting protein 3 like	chr8:26390413-26413127;chr8:26383054-26413127	Transporters/Transporter channels and pores	NM_004331
SLC2A3	solute carrier family 2 member 3	chr12:7919230-7936187	Transporters/Electrochemical Potential-driven transporters	NM_006931
PLIN2	perilipin 2	chr9:19115761-19127606;chr9:19115761-19127492	N/A	NM_001122
KDM3A	lysine demethylase 3A	chr2:86441461-86492716;chr2:86441371-86492716	N/A	NM_001146688
INSIG2	insulin induced gene 2	chr2:118088418-118110031;chr2:118088550-118110031;chr2:118088471-118110997	N/A	NM_001321329
EGLN3	egl-9 family hypoxia inducible factor 3	chr14:33924227-33951074	Enzymes/ENZYME proteins/Oxidoreductases	NM_022073
GDF15	growth differentiation factor 15	chr19:18386158-18389176	N/A	NM_004864
PTPRH	protein tyrosine phosphatase receptor type H	chr19:55181247-55209501	Enzymes/ENZYME proteins/Hydrolases	NM_002842
HCAR3	hydroxycarboxylic acid receptor 3	chr12:122714756-122716811	G-protein coupled receptors/GPCRs excl olfactory receptors	NM_006018
SPAG4	sperm associated antigen 4	chr20:35615884-35621094;chr20:35615829-35621094	N/A	NM_001317931
WSB1	WD repeat and SOCS box containing 1	chr17:27294114-27315926	N/A	NM_001348350
TMEM45A	transmembrane protein 45A	chr3:100492619-100577444	N/A	NM_018004
ADM	adrenomedullin	chr11:10305073-10307397	N/A	NM_001124
BNIP3	BCL2 interacting protein 3	chr10:131967683-131982013	Transporters/Transporter channels and pores	NM_004052
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	chr1:58780791-58784047	Transcription factors/Basic domains	NM_002228
SMAD3	SMAD family member 3	chr15:67065602-67195167;chr15:67166155-67195195;chr15:67138021-67195195;chr15:67125716-67195195	Transcription factors	NM_005902
HK2	hexokinase 2	chr2:74834127-74893358;chr2:74835170-74893358	Enzymes/ENZYME proteins/Transferases	NM_000189
GYS1	glycogen synthase 1	chr19:48968130-48993309	Enzymes/ENZYME proteins/Transferases	NM_002103

Supplementary Table 2. High frequency DE-mRNA list extracted from 10 HCC cohorts.

Supplementary Table 3. Survival analysis for high frequency DE-mRNAs in TCGA-LIHC and GSE14520.

Supplementary Table 4. Biological Processes and pathway functional enrichment analysis for high frequency DE-mRNAs.

Supplementary Table 5. Annotation for MCODE of protein-protein interaction enrichment network for high frequency DE-mRNAs.

Network	Annotation
Input ID	GO:0032787|monocarboxylic acid metabolic process|-24.6;GO:0009611|response to wounding|-23.6;hsa04610|Complement and coagulation cascades|-22.6
Input ID_MCODE_ALL	GO:0032787|monocarboxylic acid metabolic process|-16.6;hsa01230|Biosynthesis of amino acids|-16.0;GO:0044282|small molecule catabolic process|-15.0
Input ID_SUB2_MCODE_1	R-HSA-211897|Cytochrome P450 - arranged by substrate type|-23.3;R-HSA-211945|Phase I - Functionalization of compounds|-21.4;R-HSA-211859|Biological oxidations|-18.4
Input ID_SUB1_MCODE_2	R-HSA-381426|Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Bi|-5.1;R-HSA-71387|Metabolism of carbohydrates|-4.0;GO:0031589|cell-substrate adhesion|-3.8
Input ID_SUB1_MCODE_3	R-HSA-375276|Peptide ligand-binding receptors|-14.6;R-HSA-373076|Class A/1 (Rhodopsin-like receptors)|-13.0;R-HSA-418594|G alpha (i) signalling events|-12.4
Input ID_SUB1_MCODE_4	hsa01230|Biosynthesis of amino acids|-8.5;GO:0044282|small molecule catabolic process|-7.4;hsa00270|Cysteine and methionine metabolism|-6.7
Input ID_SUB1_MCODE_5	GO:0046364|monosaccharide biosynthetic process|-14.4;hsa01200|Carbon metabolism|-14.0;GO:0016051|carbohydrate biosynthetic process|-12.3
Input ID_SUB1_MCODE_6	R-HSA-5686938|Regulation of TLR by endogenous ligand|-8.8;M264|PID TOLL ENDOGENOUS PATHWAY|-8.4;R-HSA-6803157|Antimicrobial peptides|-6.6
Input ID_SUB1_MCODE_7	M166|PID ATF2 PATHWAY|-7.9;GO:0009991|response to extracellular stimulus|-5.0;R-HSA-2262752|Cellular responses to stress|-4.8

Supplementary Table 6. Survival analysis for Differential expressed miRNAs in TCGA-LIHC.

ID	log-rank P value	factor	HR	up95	low95
hsa-miR-101-3p	0.002255	protect	0.578932	0.820239	0.408615
hsa-miR-139-5p	2.43E-07	protect	0.394906	0.56083	0.278072
hsa-miR-511-5p	0.391565	protect	1.163998	1.648429	0.821928
hsa-let-7c-3p	0.509074	protect	0.88955	1.259684	0.628172
hsa-miR-125b-2-3p	0.061228	protect	0.718071	1.017179	0.506917
hsa-miR-34c-5p	0.52498	risk	1.11925	1.584939	0.79039
hsa-miR-378c	0.277779	protect	0.82486	1.168134	0.582463
hsa-miR-5589-5p	0.229327	protect	0.809008	1.14691	0.570659
hsa-miR-99a-5p	0.008725	protect	0.627161	0.888743	0.442571
hsa-miR-34c-3p	0.978891	protect	1.004682	1.422972	0.709351
hsa-miR-5589-3p	0.420898	protect	0.867101	1.228034	0.61225
hsa-miR-877-5p	0.016982	risk	1.529861	2.167074	1.080017
hsa-miR-671-5p	0.295045	risk	1.202313	1.702882	0.848888
hsa-miR-378a-3p	0.196099	risk	1.257156	1.781223	0.887279
hsa-miR-301a-3p	0.008197	risk	1.595768	2.262316	1.125605
hsa-miR-187-3p	0.538865	risk	1.11492	1.579794	0.786841
hsa-miR-138-5p	0.650067	risk	0.921018	1.311126	0.646981
hsa-miR-4524a-3p	0.399592	protect	0.861413	1.219765	0.60834
hsa-miR-223-3p	0.951984	risk	1.010698	1.431194	0.713747
hsa-let-7a-2-3p	0.679079	protect	0.929348	1.315997	0.656299
hsa-miR-375-3p	0.456514	protect	0.876533	1.2419	0.618658
hsa-miR-223-5p	0.742728	protect	0.943599	1.336783	0.666061
hsa-miR-5588-3p	0.012826	protect	0.640822	0.907489	0.452515
hsa-miR-3189-3p	0.023056	risk	1.490708	2.122278	1.047088
hsa-miR-3680-3p	0.127824	risk	1.30919	1.855126	0.923915
hsa-miR-17-3p	0.083783	risk	0.733445	1.038702	0.517898
hsa-miR-194-3p	0.165805	protect	0.782338	1.108308	0.55224
hsa-miR-135b-5p	0.072067	risk	1.376916	1.949932	0.972289
hsa-miR-7112-3p	0.515253	risk	1.123438	1.602959	0.787365
hsa-miR-205-5p	0.171105	risk	1.272456	1.804744	0.89716
hsa-miR-100-5p	0.003693	protect	0.598643	0.849333	0.421947
hsa-miR-146a-5p	0.745306	protect	1.059352	1.500225	0.748039
hsa-miR-24-2-5p	0.081881	risk	1.362826	1.9299	0.962379
hsa-miR-885-5p	0.169666	protect	0.784611	1.112856	0.553184
hsa-miR-885-3p	0.603559	protect	1.09687	1.554033	0.774195
hsa-miR-1262	0.014415	risk	1.539902	2.184093	1.085713
hsa-miR-582-5p	0.388832	risk	1.164231	1.650574	0.821189
hsa-miR-582-3p	0.022381	risk	1.493008	2.120278	1.051312
hsa-miR-210-5p	0.024344	risk	1.491626	2.113712	1.052627
hsa-miR-217-5p	0.826759	protect	1.039643	1.472817	0.733871
hsa-miR-625-3p	0.578341	protect	1.103574	1.562673	0.779354
hsa-miR-29b-1-5p	0.870187	risk	1.029333	1.457906	0.726746
hsa-miR-21-3p	0.060608	risk	1.393654	1.97479	0.983533
hsa-miR-561-5p	0.0075	risk	1.62099	2.40279	1.093566
hsa-miR-541-3p	0.568232	protect	1.106522	1.566843	0.781438
hsa-miR-216b-3p	0.46815	risk	1.150479	1.701683	0.77782
hsa-miR-1224-5p	0.886944	risk	1.025507	1.452126	0.724225
hsa-miR-5588-5p	0.089422	protect	0.739409	1.047195	0.522086
hsa-miR-2355-3p	0.008464	risk	1.595746	2.261333	1.126063
hsa-miR-216a-3p	0.335209	risk	1.186604	1.680295	0.837965
hsa-miR-4449	0.844205	protect	1.035457	1.466219	0.731249
hsa-miR-653-5p	0.577505	protect	1.103923	1.563284	0.779543

hsa-miR-194-5p	0.004083	protect	0.601321	0.851918	0.424439
hsa-miR-548v	0.174238	risk	1.271901	1.802342	0.897572
hsa-miR-141-3p	0.535513	risk	1.115914	1.580405	0.787941
hsa-miR-216a-5p	0.342731	protect	1.18372	1.676152	0.835959
hsa-miR-155-5p	0.45926	risk	1.140173	1.614518	0.805191
hsa-miR-216b-5p	0.131044	protect	1.308972	1.853506	0.924414
hsa-miR-155-3p	0.832344	risk	1.038098	1.470129	0.733029
hsa-miR-210-3p	0.000975	risk	1.791011	2.541255	1.262258
hsa-miR-629-5p	0.040936	protect	1.442479	2.042647	1.018651
hsa-miR-642a-5p	0.690464	risk	1.073046	1.519633	0.757701
hsa-miR-200c-3p	0.038228	risk	1.443662	2.045728	1.018786

Supplementary Table 7. Differential expressed lncRNAs between high hypoxia score and low hypoxia score groups in TCGA-LIHC.

Supplementary Table 8. Survival analysis for differential expressed lncRNAs in TCGA-LIHC.

Supplementary Table 9. Significant difference in CNAs events between high hypoxia score and low hypoxia score groups in TCGA-LIHC.

Supplementary Table 10. Significant difference in DNA methylation level between high hypoxia score and low hypoxia score groups in TCGA-LIHC.

Supplementary Table 11. DNA methylation changes of 30 up-regulated genes in high hypoxia score groups.

site	Gene	DNA methylation level	Position_to_TSS
cg23174662	HIF1A	reduced	3587;3824;3824;1742;3851;469;3851;1742
cg04293307	AXIN2	reduced	4185;1267;2864;4185;83571
cg06075789	ZFP36	reduced	.
cg17382541	MYADM	reduced	1523;1227;-299;-299;1361;-299;-299
cg09442403	TM4SF1	reduced	6355;4202;6182;6232
cg03707168	PPP1R15A	reduced	3477;1551
cg04797170	TMSB10	reduced	252;252
cg12232146	PHLDA1	reduced	29859
cg12457415	CLIC1	reduced	1488;2234;1483;4679;1461
cg26403843	RNF145	reduced	750;2475;2977;997;2715;558;1056;2522
cg05374271	RGS2	reduced	718;717;720;616
cg27270412	NRP1	reduced	2859;1119;1119;1239;979;1502;1233;1502;1675
cg16643088	CSRNP1	reduced	6324;7311
cg13275603	ENC1	reduced	9763;601;9763;8988;9763
cg22502206	ANXA2	reduced	-641;-640;-640;-583;4257;-640;-640;-653;-640;-638;-652;-640;-593;-662;-640;-640;-648;-638;-625;-652;116;-642;-658;-1288;-640;-640;-638;-640;-640;-640;-640
cg09055236	TTYH3	reduced	1611;1406
cg05132999	TTC39A	reduced	-1141;32;-752;-1112;13802;-367;13802;-750;13802;1660
cg10523105	PNMA1	reduced	2100
cg15460348	SPP1	reduced	89;42;68;42;-1124;39;42;42;41;42;106
cg03359362	SLC1A5	reduced	-1462;-1462;989;2241;-1448
cg05117638	MARCKS	reduced	18616;1056
cg11230435	MAP3K8	reduced	1507;1200;779;1303;1423
cg27505627	BAIAP2L2	reduced	64;89
cg00937742	LYPD1	reduced	-146
cg22679626	TMSB4X	reduced	.
cg19913563	IER3	reduced	.
cg22249612	CD63	reduced	1520;950;1960;200;-1344;-641;-592;2007;1300;115;947;-572;1277;782;649;-1404
cg04860291	TKT	reduced	3379;3388;3366;3388;3397;3340;3353;3418
cg12586150	SERPINB1	reduced	1449;1305;1291;1402
cg24147596	ARL14	reduced	472

Supplementary Table 12. Significant difference in RNA alternative splicing events between high hypoxia score and low hypoxia score groups in TCGA-LIHC.