# Comprehensive characterization of enhancer RNA in hepatocellular carcinoma reveals three immune subtypes with implications for immunotherapy

Xiaoyun Bu,[1,2,3] Shuang Liu,[1,2,3] Dongsheng Wen,[1,2,3] Anna Kan,[1,2] Yujie Xu,[1,2] Xuanjia Lin,[1] and Ming Shi[1,2]

[1]Department of Liver Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China; [2]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

**Hepatocellular carcinoma (HCC) is highly heterogeneous. Molecular subtyping for guiding immunotherapy is warranted. Previous studies have indicated that enhancer RNAs (eRNAs) are involved in tumor heterogeneity and immune infiltration. However, the eRNA landscape and its correlation with immune infiltration in HCC remain unknown. Here we first revealed the genome-wide eRNA landscape in two HCC cohorts. Then we divided individuals with HCC into three immune-related clusters (C1, C2, and C3) based on eRNA expression profiles. The prognosis, biological properties, immune infiltration, clinical features, genomic features, and drug response were analyzed. C1 was enriched in immune infiltration and potentially sensitive to immune checkpoint inhibitors (ICIs). C2 displayed features of immune depletion, high proliferation activity, malignant clinical features, and the worst prognosis. C2 may benefit from targeted therapy. C3 presented moderate immune infiltration, metabolism-related signatures, and the best prognosis. Transarterial chemoembolization (TACE) may be effective for C3. Finally, we constructed a 51-eRNA classifier for subtype prediction and validated its efficacy in The Cancer Genome Atlas (TCGA) cohort and Sun Yat-sen University Cancer Center (SYSUCC) cohort. Our results provide a novel method for immune classification of HCC, shed new light on tumor heterogeneity, and may aid in HCC immunotherapy.**

## INTRODUCTION

Liver cancer is the second leading cause of cancer-related death worldwide.[1] Hepatocellular carcinoma (HCC) is the most prevalent pathological type of liver cancer. Although several recent improvements have been made in comprehensive therapy of HCC, the 5-year survival of individuals with HCC still remains poor.[2] HCC is a highly heterogeneous disease.[3] Molecular subtyping has important implications when devising individualized therapeutic strategies for individuals with HCC.[4]

Enhancers are essential regulatory elements of DNA that activate target gene transcription.[5,6] Tens of thousands of enhancers have been identified in the Encyclopedia of DNA Elements (ENCODE).[7] Although enhancers generally regulate gene expression in *cis*, they also exert long-range effects, which are independent of position and orientation when activated by chromosomal rearrangement or specific transcription factors.[8–10] Recently, with the development of RNA sequencing technology, it has been reported that enhancers also function as noncoding RNA transcription units to generate enhancer RNAs (eRNAs), which can be used to infer the activities of enhancer loci.[11] As an important component of the transcriptome and epigenome, accumulating evidence shows that eRNAs are widely involved in human cancer by modulating chromatin structure or interacting with transcriptional regulators.[12–14] Qin et al.[15] performed molecular subtyping of lung adenocarcinoma based on functional eRNAs, which identified distinct immune signatures among different subtypes and might be instructive for cancer immunotherapy. However, the genome-wide eRNA landscape and its correlation with the immune microenvironment in HCC remain largely undefined.

In this study, we illustrated the genome-wide landscape of eRNAs in HCC by analyzing total RNA sequencing data from the GEO database. We established a novel immune-related HCC subtyping strategy based on the expression profile of eRNAs, and the prognosis, immune infiltration, biological properties, clinical features, genomic features and drug responses of different subtypes were analyzed. Finally, we generated a 51-eRNA classifier for subtype prediction. The efficacy of the classifier was validated in HCC cohorts from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) and Sun Yat-sen University Cancer Center (SYSUCC).

## RESULTS

### Landscape of eRNAs in HCC

The main research processes of this study are summarized in Figure 1. We first explored the genomic distribution of eRNAs and found that they were relatively uniform across all chromosomes except chromosome 13 (Figure 2A). The length of most
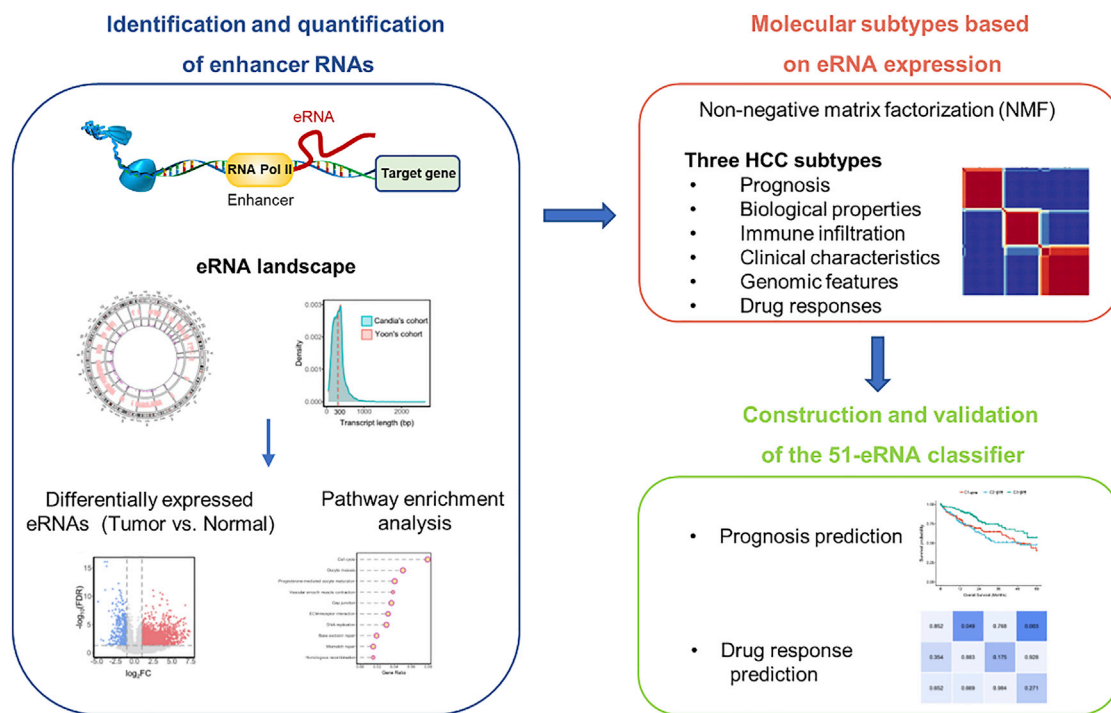
**Figure 1. Overview of the study workflow**

eRNAs was around 300 bp in two cohorts (Figure 2B). The average expression level of global eRNA was significantly upregulated in HCC compared with that in normal liver (Figure 2C), implying that eRNAs may be involved in HCC carcinogenesis. Next we analyzed the average expression level of global eRNA in different groups based on age, sex, α-fetoprotein (AFP) level, tumor size, Tumor Node Metastasis (TNM) stage and hepatitis viral infection status. No significant difference in the average expression level of global eRNA was observed between different age, sex, AFP level, and tumor size groups (Figures S1A–S1D), except for remarkably downregulated expression of global eRNAs in hepatitis C virus (HCV)-infected samples versus hepatitis B virus (HBV)-infected samples in the Yoon et al.[17] cohort (Figure S1F), which suggested a distinct regulation pattern because of different virus infections. eRNAs tended to be highly expressed in advanced HCC tissues in the Candia et al.[16] cohort, although it was not statistically significant (p = 0.28; Figure S1E).

Differentially expressed eRNAs (DEeRNAs) in HCC and normal liver were identified (Figure 2D). There were 68 upregulated eRNAs and 55 downregulated eRNAs in both cohorts (Figure 2F). The genomic distribution of DEeRNAs is presented in Figure 2E. Functional enrichment analyses of co-expressed protein-coding genes (PCGs) of DEeRNAs were performed. We focused on eRNAs upregulated in HCC and found that PCGs correlated to these eRNAs were mainly involved in cell cycle, DNA replication, mismatch repair, and several metabolism-related pathways, indicating that these eRNAs may play essential roles in HCC carcinogenesis (Figure 2G; Table S2).

## Identification of three HCC subtypes based on expression profiles of eRNAs

To explore HCC molecular subtyping based on expression profiles of eRNAs, we conducted non-negative matrix factorization (NMF) analyses in two training cohorts (the Candia et al.[16] and Yoon et al.[17] cohorts). The cophenetic correlation coefficients were calculated to decide the appropriate number of subtypes, and k = 3 was chosen (Figures 3A and 3D). We classified individuals with HCC into three subtypes: clusters 1, 2, and 3 (C1, C2, and C3, respectively) (Figures 3B and 3E). Principal-component analysis (PCA) showed that individuals in the three subtypes were well separated (Figures 3C and 3F). SubMap analyses confirmed that C1, C2, and C3 in the Candia et al.[16] cohort were highly associated with corresponding subtypes in the Yoon et al.[17] cohort (Figure 3G). Significant differences between subtypes were observed in overall survival (OS) or disease-free survival (DFS). C3 had a survival advantage over C1 and C2 (OS in the Candia et al.[16] cohort, C3 versus C1, p = 0.029, C3 versus C2, p < 0.001; DFS in the Yoon et al.[17] cohort, C3 versus C1, p = 0.025, C3 versus C2, p < 0.001) (Figures 3H and 3I). The subtype-specific eRNAs are shown in Figures S2A and S2B and Table S4.

## Prognostic significance of the eRNA subtyping system in HCC

We first evaluated the prognostic significance of the eRNA subtyping system in the Candia et al.[16] and Yoon et al.[17] cohorts. Univariate and multivariate Cox regression identified C3 as an independent prognostic factor in the Candia et al.[16] and Yoon et al.[17] cohorts (Figures 4A and 4D). Then we evaluated the efficacy of the eRNA subtyping system in prognosis prediction through time-dependent
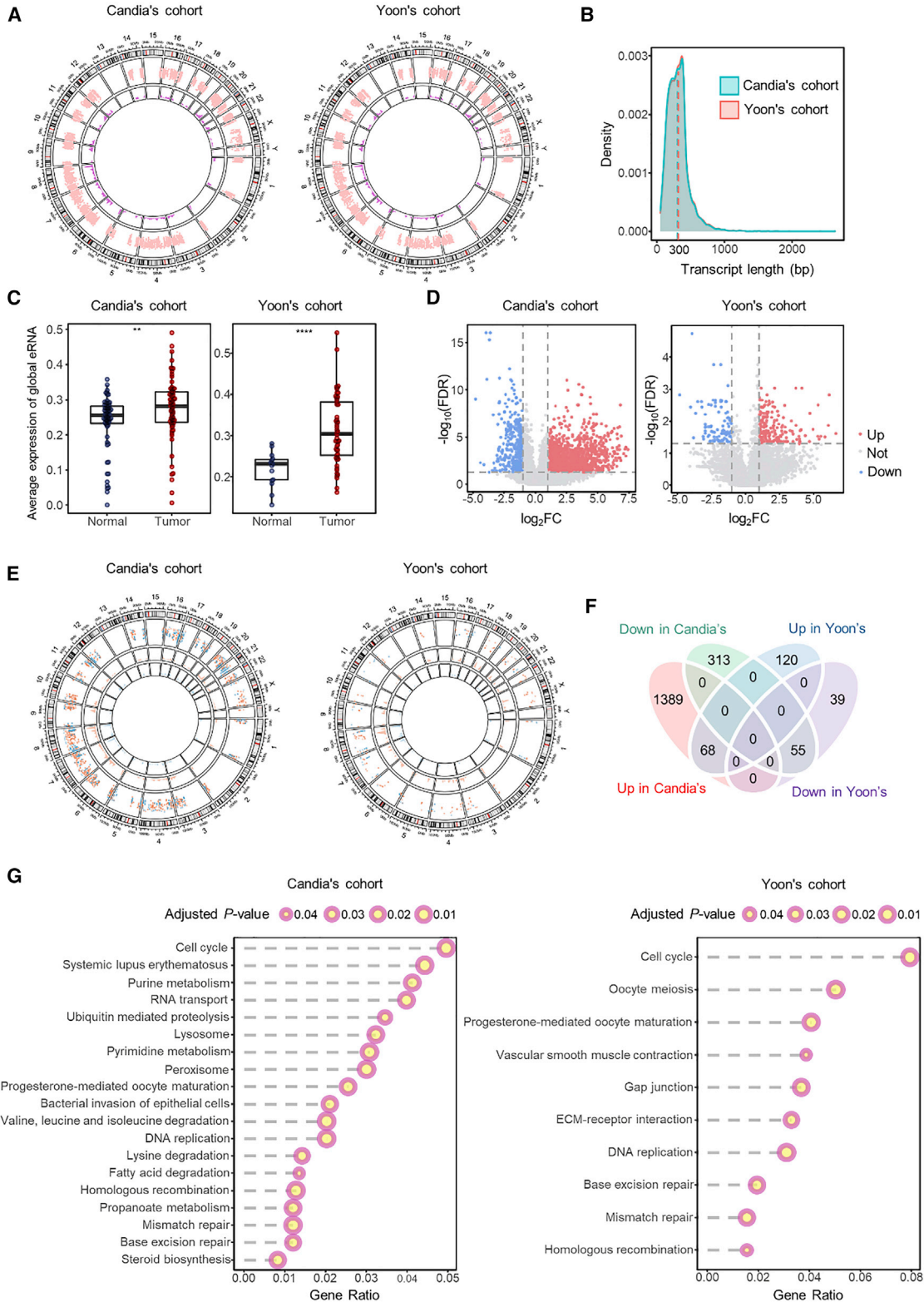
*(legend on next page)*

receiver operating characteristic (ROC) analyses. The area under the ROC curve (AUC) for 1-, 3- and 5-year OS was 0.8, 0.71, and 0.64 in the Candia et al.[16] cohort and 0.71, 0.72, and 0.79 for 1-, 2-, and 3-year DFS in the Yoon et al.[17] cohort (Figures 4B and 4E). We also compared the predictive ability of eRNA subtype with other clinical parameters. The results showed that the predictive ability of the eRNA subtyping system was superior to the predictive ability of TNM staging system, serum AFP level, tumor size, or cirrhosis (Figures 4C and 4F).

### Biological properties of the three HCC subtypes

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and cancer hallmark gene set enrichment analyses were utilized to distinguish the characteristic pathways related to the three subtypes. As shown in Figures 5A and 5B, pathways involved in metastasis, immune infiltration, and tumorigenesis were significantly enriched in C1, indicating that C1 was correlated to immune infiltration. C2 was inactive in immune-related pathways but had a remarkable enrichment in proliferation-related pathways, including G2M checkpoint, DNA replication, mismatch repair, E2F targets, and cell cycle. C3 presented moderate activity in immune-related pathways and was activated in metabolism-related pathways, including adipogenesis, bile acid metabolism, and drug metabolism, suggesting that C3 was a well-differentiated subtype of HCC. Generally, C1 had a more activated immune microenvironment compared with C2 and C3.

### Characterization of immune infiltration in the three HCC subtypes

Pathway enrichment analyses revealed that the extent of immune infiltration was significantly different in the three subtypes. We analyzed the immune and stromal scores according to the estimation of stromal and immune cells in malignant tumours using expression data (ESTIMATE) algorithm (Figures 6A and 6B). C1 displayed markedly higher immune and stromal scores than C2 and C3 in the Candia et al.[16] cohort. A similar trend was observed in the Yoon et al.[17] cohort, which was in line with the results of functional enrichment analyses. C1 also exhibited higher cytolytic activity scores than C2 and C3, indicating activation of the tumor microenvironment (TME) in C1 (Figure 6C).

We also evaluated the expression level of 15 immune checkpoint genes that may become potential therapeutic targets for HCC in the three subtypes. C1 showed the highest expression of 13 genes (except for LAG3 and OX40L) among the three subtypes, suggesting that individuals in C1 may benefit from immune checkpoint inhibitors (ICIs). Both cohorts exhibited similar results (Figure 6D).

Then we analyzed immune infiltration cells in the three HCC subtypes through tumor immune estimation resource (TIMER), microenvironment cell populations (MCP)-counter, and xCell. C1 showed significant enrichment of immune cells, including B cells, CD8+ T cells, and dendritic cells. C1 and C3 displayed higher enrichment of endothelial cells and cancer-associated fibroblast cells. Consistent with the immune and stromal scores, C2 showed the lowest activity in immune cells and stromal cells (Figure 6E). Our eRNA subtyping system distinguished three immune-related HCC subtypes. C1 was defined as an immune-enriched subtype and C2 as an immune-depleted subtype. C3 was moderately infiltrated by immune cells.

### Clinical characteristics of the three HCC subtypes

We next investigated clinical features of the three subtypes in the Candia et al.[16] and Yoon et al.[17] cohorts. In the Candia et al.[16] cohort, C2 had a higher level of AFP (Figure 7A). Similarly, C2 also had more aggressive clinical features in the Yoon et al.[17] cohort, such as larger tumor size and increased relapse rate (Figure 7B). These results confirmed that C2 was correlated with poor prognosis. Other clinical characteristics, such as age, sex, virus infection, tumor numbers, and Child-Pugh score, showed no significant difference among the three subtypes.

Considering that several HCC classifications have been reported so far, we estimated the correlation between eRNA subtype and previous classifications, including the Boyault et al.[18] subtype, Hoshida et al.[26] subtype, Lee et al.[24] subtype, and Roessler et al. subtype (Figures 7C and 7D). In the Candia et al.[16] cohort, the C1 subtype was significantly linked to the Boyault et al.[18] G4 class, the Hoshida et al. S1 class, the Lee et al. A class, and the Roessler et al.[19] high-risk group (p < 0.001). The C2 subtype was strongly correlated with the Boyault et al.[18] G2/3 class, the Hoshida et al. S2 class, the Lee et al. A class, and the Roessler et al.[19] high-risk group (p < 0.001). The C3 subtype was associated with the Boyault et al.[18] G4/6 class, the Hoshida et al. S3 class, the Lee et al. B class, and the Roessler et al.[19] low-risk group (p < 0.001). Similarly, in the Yoon et al.[17] cohort, the C1 subtype was significantly correlated with the Boyault et al.[18] G4 class, the Hoshida et al. S1 class, the Lee et al. A class, and the Roessler et al.[19] high-risk group (p < 0.001). The C2 subtype was strongly linked to the Boyault et al.[18] G4/5 class, the Hoshida et al. S2 class, and the Roessler et al.[19] high-risk group (p < 0.001). The C3 subtype was associated with the Boyault et al.[18] G4 class, the Hoshida et al. S3 class, the Lee et al. B class, and the Roessler et al. low-risk group (p < 0.001). These results revealed that eRNA subtype had great consistency with previous studies.

**Figure 2. Landscape of eRNA in HCC**

(A) The genomic distribution of eRNA in the Candia et al.[16] (left) and Yoon et al.[17] cohorts (right). Pink dots, locations; purple peaks, density of eRNA distribution. (B) Length distribution of eRNAs in two cohorts. (C) Average expression of global eRNAs in HCC and normal liver tissues in two cohorts. (D) Volcano plot of DEeRNAs in HCC and normal liver tissue in two cohorts. Red, upregulated; white, not significantly different; blue, downregulated. (E) The distribution of DEeRNAs on 23 chromosomes in two cohorts. Orange dots, location of upregulated eRNAs; blue dots, location of downregulated eRNAs; peaks, density of eRNA distribution. (F) Venn diagram of DEeRNAs in two cohorts. (G) Functional enrichment analysis of upregulated eRNAs in two cohorts. Boxplot "boxes" indicate the first, second, and third quartiles of the data.*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
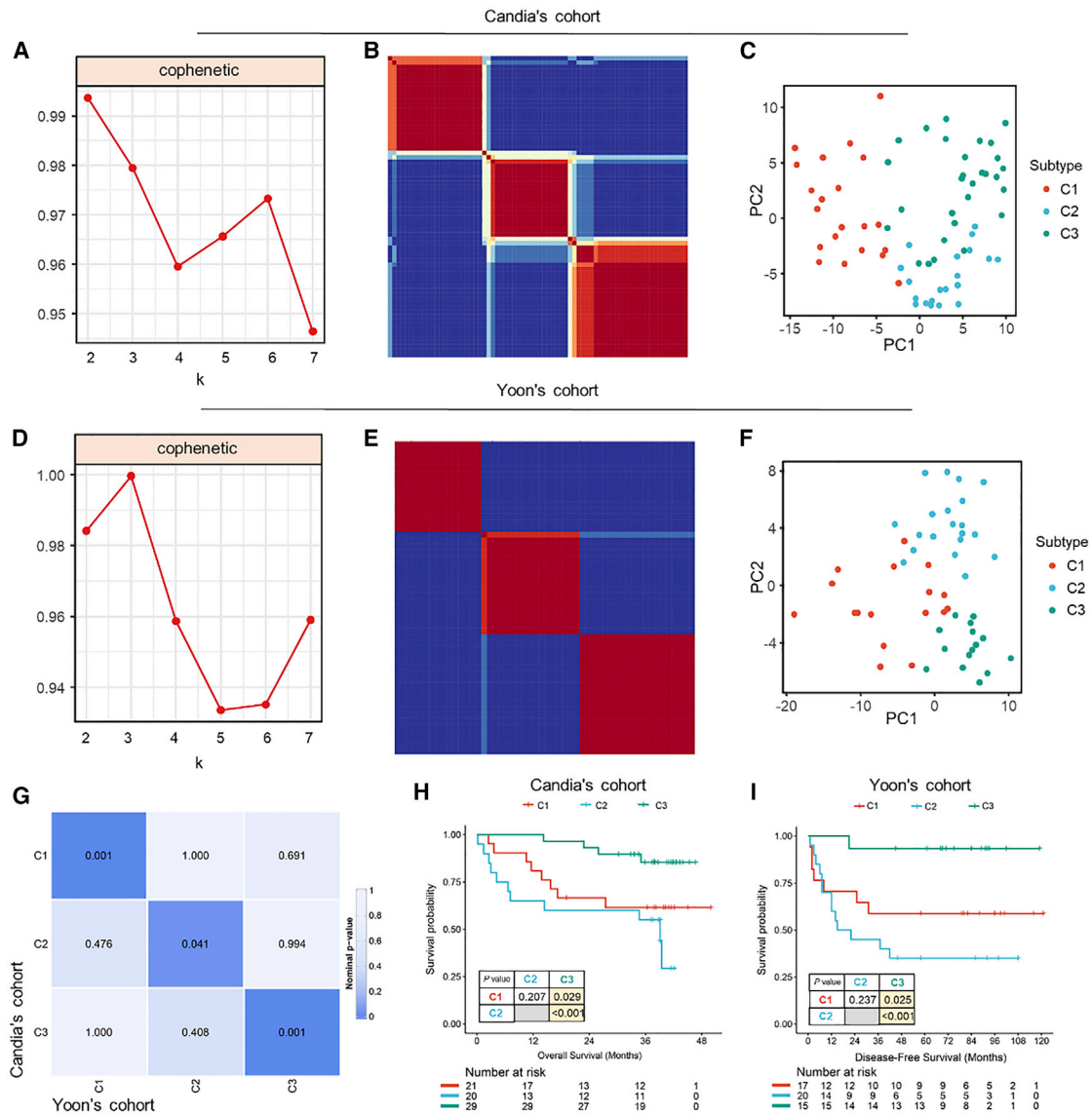
**Figure 3. Identification of three HCC subtypes based on expression profiles of eRNAs**

(A) Cophenetic correlation coefficient for k = 2–7 for the Candia et al.[16] cohort. (B) NMF subtyping based on eRNA expression profile in the Candia et al.[16] cohort. (C) PCA of eRNAs to divide individuals with HCC into three subtypes in the Candia et al.[16] cohort. (D) Cophenetic correlation coefficient for k = 2–7 for the Yoon et al.[17] cohort. (E) NMF subtyping based on eRNA expression profile in the Yoon et al.[17] cohort. (F) PCA of eRNAs to divide individuals with HCC into three subtypes in the Yoon et al.[17] cohort. (G) SubMap analysis of GEPs between the HCC subtypes in the Candia et al.[16] and Yoon et al.[17] cohorts. (H) OS of three subtypes in the Candia et al.[16] cohort. (I) DFS of three subtypes in the Yoon et al.[17] cohort.

## Genomic features of different immune subtypes

To investigate whether gene somatic mutation was distinct among the three subtypes, we focused on genes involved in cell cycle regulation, Wnt/β-catenin signaling, hepatic differentiation, and chromatin modification in the Candia et al.[16] cohort (Figures 8A and 8B). Generally, the C1 subtype had the lowest gene mutation burden among the three subgroups (Figure 8C). The mutation frequency of cell-cycle-related genes was higher in C2 (53%) than C1 (22%) and C3 (46%), which was consisted with the proliferative feature of C2. Although

the mutation frequency of cell-cycle-related genes was also common in C3, the function enrichment result and the correlation analysis with other HCC classifications did not support C3 as a high-proliferation subtype. Albumin (ALB) and apolipoprotein B (APOB) mutation were most common in the C3 subtype (50%) compared with C1 (5.5%) and C2 (21%), in agreement with the dysregulation of cellular metabolism in C3. CTNNB1 mutation was most frequent in the C3 subtype (39.3%), suggesting that the Wnt/β-catenin pathway may be dysregulated in C3. No significant difference in
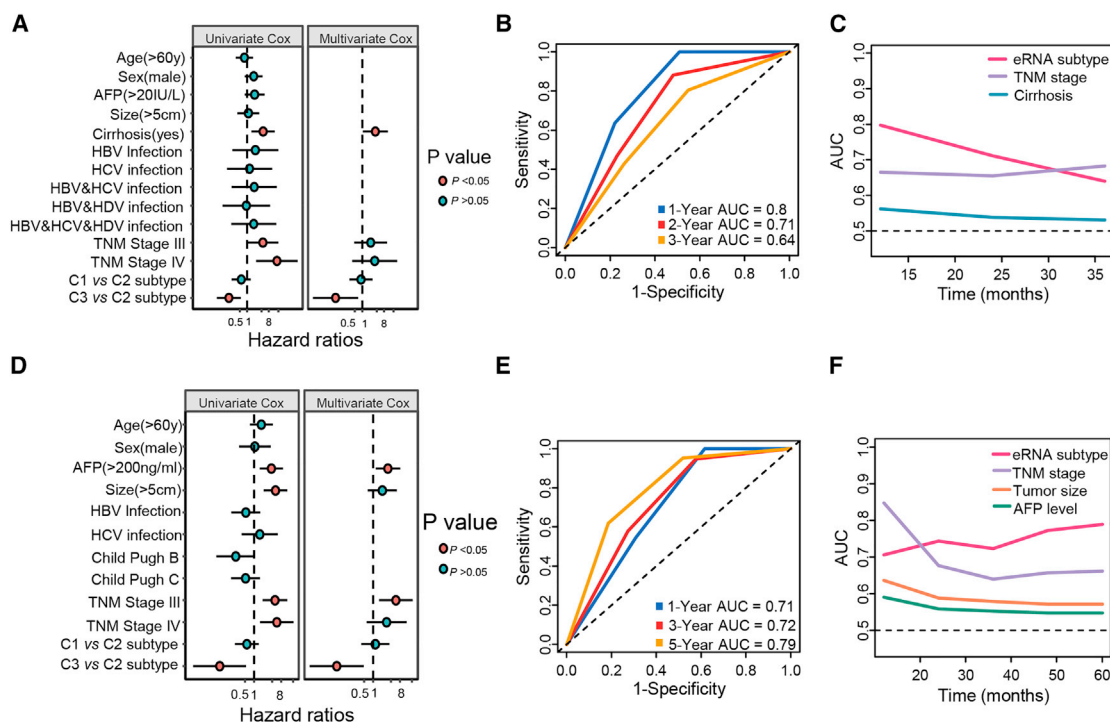
**Figure 4. Prognostic significance of the eRNA subtyping system in HCC**

(A) Univariate and multivariate Cox regression analyses of OS in the Candia et al.[16] cohort. (B) Time-dependent ROC analysis of eRNA subtyping in the Candia et al.[16] cohorts. (C) The sensitivity and specificity of eRNA subtyping, TNM stage, and cirrhosis in prognostic prediction were estimated by time-dependent ROC curves in the Candia et al.[16] cohort. (D) Univariate and multivariate Cox regression analyses of DFS in the Yoon et al.[17] cohort. (E) Time-dependent ROC analysis of eRNA subtyping in the Yoon et al.[17] cohort. (F) The sensitivity and specificity of eRNA subtyping, TNM stage, tumor size, and AFP level in prognostic prediction were estimated by time-dependent ROC curves in the Yoon et al.[17] cohort.

chromatin-modification-related gene mutations was observed among the three subtypes (Figure 8A).

**Potential response to targeted therapy, immunotherapy, and transarterial chemoembolization (TACE) in the three HCC subtypes**

Considering that the immune microenvironment is closely related to drug response in cancer treatment, we explored whether eRNA-based classification could be instructive for developing an individualized therapeutic strategy in clinical practice. We compared the gene expression profiles (GEPs) of the three immune subtypes with HCC cell lines that had drug sensitivity data. For targeted agents (sorafenib and cabozantinib), C2 showed a significant correlation with the sorafenib and cabozantinib response group (sorafenib, p = 0.003; cabozantinib, p = 0.015) in the Yoon et al.[17] cohort (Figure 9A). The same tendency was observed in the Candia et al.[16] cohort, but the results were not all statistically significant (sorafenib, p = 0.015; cabozantinib, p = 0.053) (Figure 9D). On the contrary, C1 was strongly associated with the sorafenib- and cabozantinib-resistant group (Candia et al.[16] cohort, sorafenib, p = 0.004, cabozantinib, p = 0.018; Yoon et al.[17] cohort, sorafenib, p = 0.002, cabozantinib, p = 0.011) (Figures 9A and 9D). These results showed that C2 was more likely to benefit from sorafenib and cabozantinib treatment.

C1 was more likely to respond to the programmed cell death protein 1 (PD-1) inhibitor in both cohorts (Candia et al.[16] cohort, p = 0.014; Yoon et al.[17] cohort, p = 0.033) (Figures 9B and 9E). The same tendency was observed for cytotoxic T lymphocyte-associated protein 4 (CTLA-4) inhibitor response, but the results were not all statistically significant (Candia et al.[16] cohort, p = 0.131; Yoon et al.[17] cohort, p = 0.033). However, C2 seemed to be immunotherapy resistant because of immune depletion features in the Candia et al.[16] cohort (PD-1 inhibitor, p = 0.008; CTLA-4 inhibitor, p = 0.045) (Figure 9B).

TACE is a widely used treatment for HCC.[20,21] We analyzed the potential TACE response status in the three HCC subtypes and found that the GEP of C3 was significantly correlated with the TACE response group (Candia et al.[16] cohort, p = 0.001; Yoon et al.[17] cohort, p = 0.004) (Figures 9C and 9F), indicating that C3 tends to benefit from TACE treatment.

**Construction and validation of a 51-eRNA classifier for eRNA subtype identification**

For subtype prediction, we performed differential expression analysis to construct a subtype-specific eRNA classifier. After integrating subtype-specific eRNAs in the Candia et al.[16] and Yoon et al.[17] cohorts, a
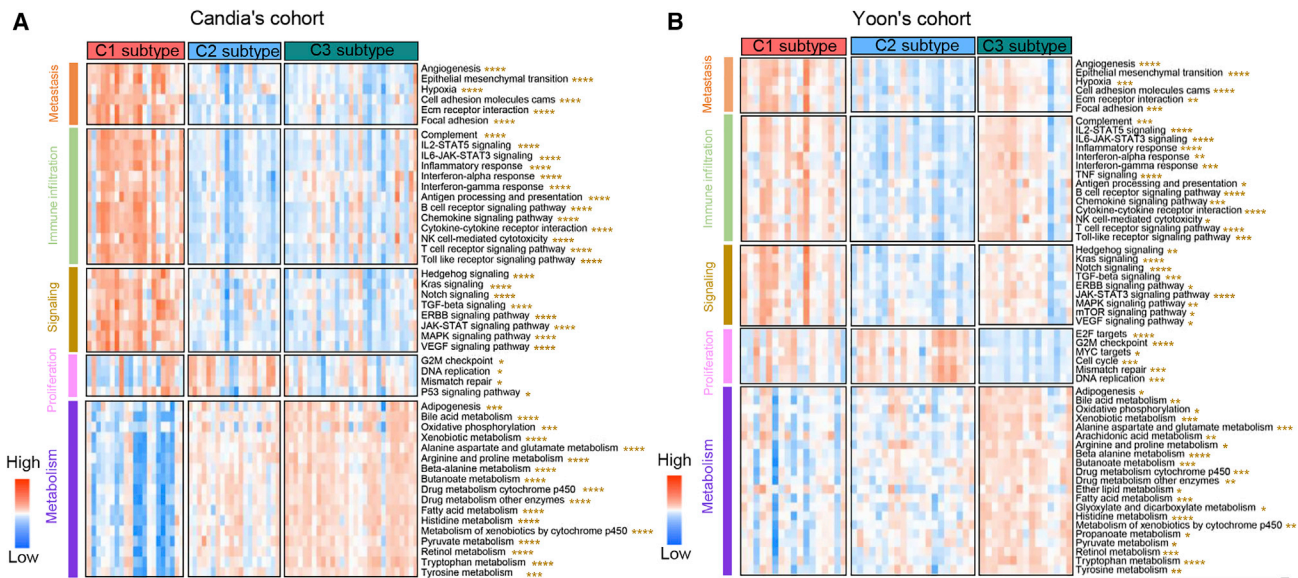
**Figure 5. Functional enrichment analyses of three HCC subtypes**

(A and B) Heatmaps of biological properties enriched in three subtypes in the Candia et al.[16] (A) and Yoon et al.[17] (B) cohorts, analyzed by ssGSEA. High and low ssGESA scores are represented in red and blue, respectively. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

total of 51 eRNAs, called the 51-eRNA classifier, were screened out (Figure S2C; Table S5). High accordance was observed between subtypes predicted by the 51-eRNA classifier based on NTP and original classification based on NMF, with concordance of 80.77% in C1, 72.72% in C2, and 95.45% in C3 in the Candia et al.[16] cohort and 88.24% in C1, 100% in C2, and 73.68% in C3 in the Yoon et al.[17] cohort (Figure 10A).

To test the efficacy of the subtype-specific eRNA classifier, we divided the samples in the two validation cohorts into three subtypes based on the 51-eRNA classifier using the NTP algorithm. The prognosis, immune infiltration, biological properties, genomic features, and drug responses among the three subgroups were analyzed. As expected, C1-predicted displayed the highest levels of immune cell infiltration and most active immune-related pathways. C2-predicted was still characterized by least immune cell infiltration, high proliferation, and poor prognosis. C2-predicted showed a higher copy number alteration (CNA) burden than the C1- and C3-predicted subtypes in the TCGA-LIHC cohort. C3-predicted presented moderate immune cell infiltration and metabolism signatures (Figures 10B–10F, S3, and S4). Similar results were observed in the SYSUCC cohort. Immunohistochemistry (IHC) staining demonstrated the highest CD8+ T cell infiltration in the C1 subtype (Figure 10G). C2 displayed the most abundant ki67-positive cells, consistent with high proliferation in C2 (Figure 10G). These results showed that our eRNA subtypes were highly reproducible and that the 51-eRNA classifier was reliable.

In accordance with the result of the Kaplan-Meier analysis (Figure 10B), univariate Cox regression revealed that C3 had a lower

survival risk than C2 in the TCGA-LIHC cohort (Figure S5A). The AUC of the 51-eRNA classifier for 1-, 3-, and 5-year OS was 0.64, 0.60, and 0.57 in the TCGA-LIHC cohort (Figure S5B). The efficacy of the eRNA subtyping system was shown to be superior to the TNM staging system and vascular invasion in prognosis evaluation (Figure S5C).

The efficacy of the classifier in prediction treatment responses was also validated in the TCGA-LIHC and SYSUCC cohorts. C1-predicted was sensitive to ICIs (TCGA-LIHC cohort, PD1 inhibitor, p = 0.003, CTLA4 inhibitor, p = 0.049 [Figure 11B]; SYSUCC cohort, PD1 inhibitor, p = 0.001 [Figure 11E]). C2-predicted still showed a significant correlation with the sorafenib- and cabozantinib-response group (TCGA-LIHC cohort, sorafenib, p = 0.001, cabozantinib, p = 0.001 [Figure 11A]; SYSUCC cohort, sorafenib, p = 0.003, cabozantinib, p = 0.040 [Figure 11D]). C3-predicted was likely to benefit from TACE (TCGA-LIHC cohort, p = 0.001 [Figure 11C]; SYSUCC cohort, p = 0.002 [Figure 11F]). The 51-eRNA classifier is highly valuable for predicting therapeutic responses of individuals with HCC.

A summary of the three immune-related HCC subtypes is shown in Figure 12.

## DISCUSSION

HCC is a highly heterogeneous disease. Several staging systems for HCC (e.g., TNM and the Barcelona clinic liver cancer staging system) are generally based on clinical information and lack a systemic overview of molecular features, which is insufficient for guiding individualized therapy, such as immunotherapy and target therapy.[22,23] Most
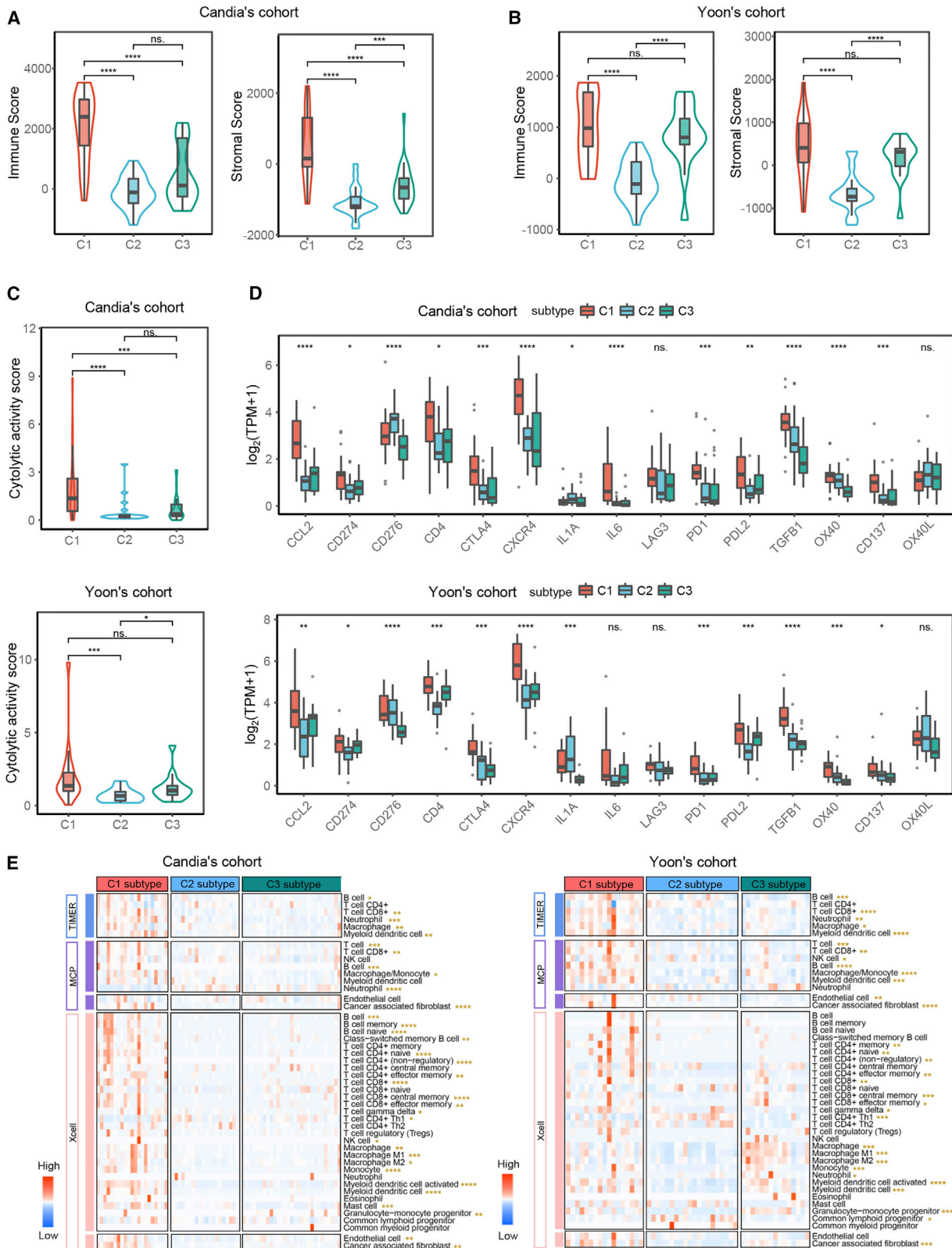
molecular models of HCC are based on gene expression and mutation.[18,19,24–26] Few studies have been performed to explore epigenetic signatures during tumorigenesis.

eRNAs are important parts of the epigenetic regulation system transcribed by activated enhancers.[27,28] Accumulated evidence indicates that eRNAs are strongly associated with cancer development.[29] In this study, we characterized the eRNA landscape of HCC and uncovered its important role in the cell cycle, extracellular matrix (ECM)-receptor interaction, DNA replication, mismatch repair, and metabolism. We constructed a novel three-subtype molecular classification system for HCC based on eRNA expression profiles. The prognosis, immune signature, clinical features, biological properties, and drug responses were distinct among the three subtypes. C1 was distinguished by enriched immune infiltration and high expression of immune-checkpoint-related genes, which implied that C1 samples were "hot" tumors and sensitive to ICIs. C2 exhibited a lack of immune infiltration, high proliferation activity, and poor prognosis. C2 might benefit from targeted therapy. C3 was defined as a moderate immune subtype with metabolism-related signatures, low AFP level, and good prognosis. TACE was predicted to be effective for C3.

The tumor immune microenvironment has been recognized as a key factor in HCC biological behavior and drug response.[30] Immune classification has the potential to predict prognosis and treatment responses for individuals with HCC.[31] In recent years, immunotherapy has achieved inspiring outcomes for individuals with HCC.[32–34] In this study, we found that C1 represented a small group of individuals with HCC who might be suitable for immunotherapy. The high immune score and enriched immune infiltration of C1 are in line with the notion that tumors with abundant immune cell infiltration are more likely to respond to immunotherapy. C1 accounts for about 30% of individuals in every HCC cohort, which is in agreement with the objective response rate of ICIs in clinical trials.[33,35,36] C2 presented deficiency in immune infiltration cells, suggesting that immunotherapy may be ineffective for C2. The enrichment of TP53 mutations and high proliferation of C2 may contribute to its resistance to ICIs.[37] Previous studies have revealed that sorafenib responders were more likely to be highly proliferative and ICI resistant, which is consist with the C2 subtype in this study.[38]
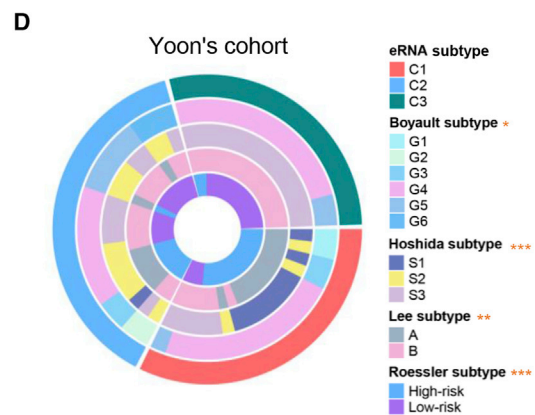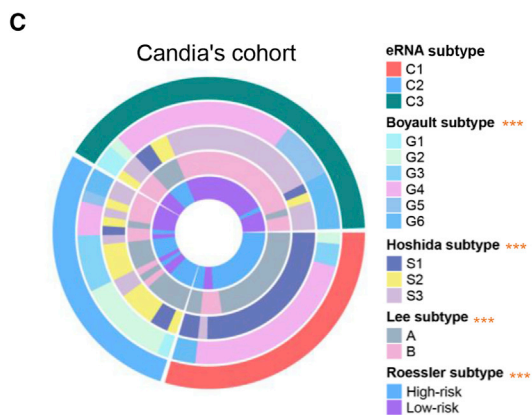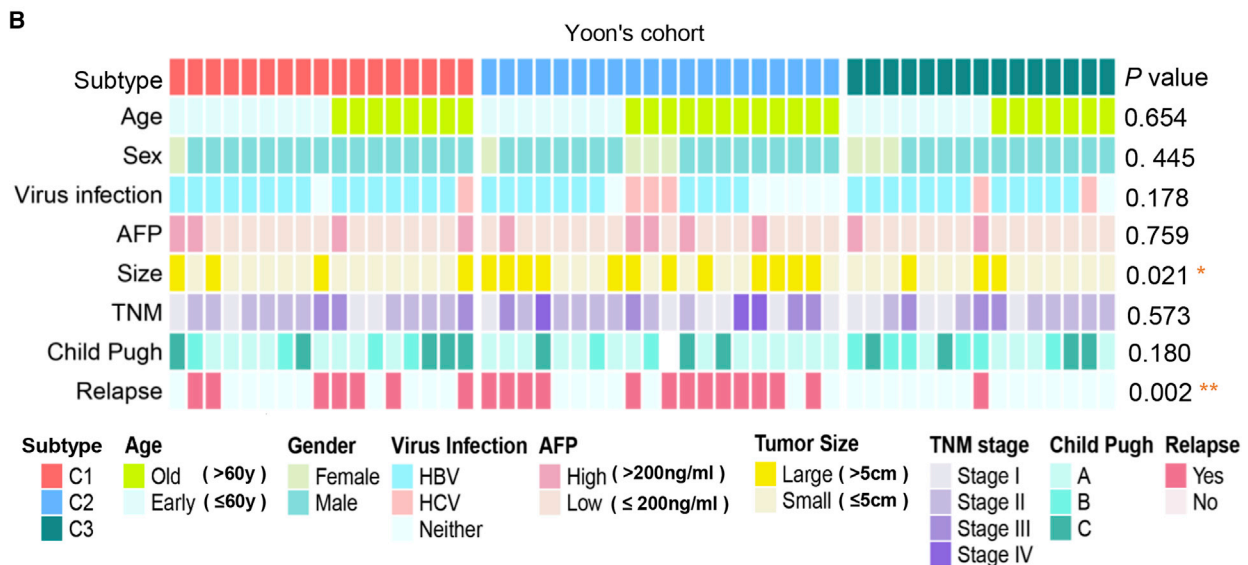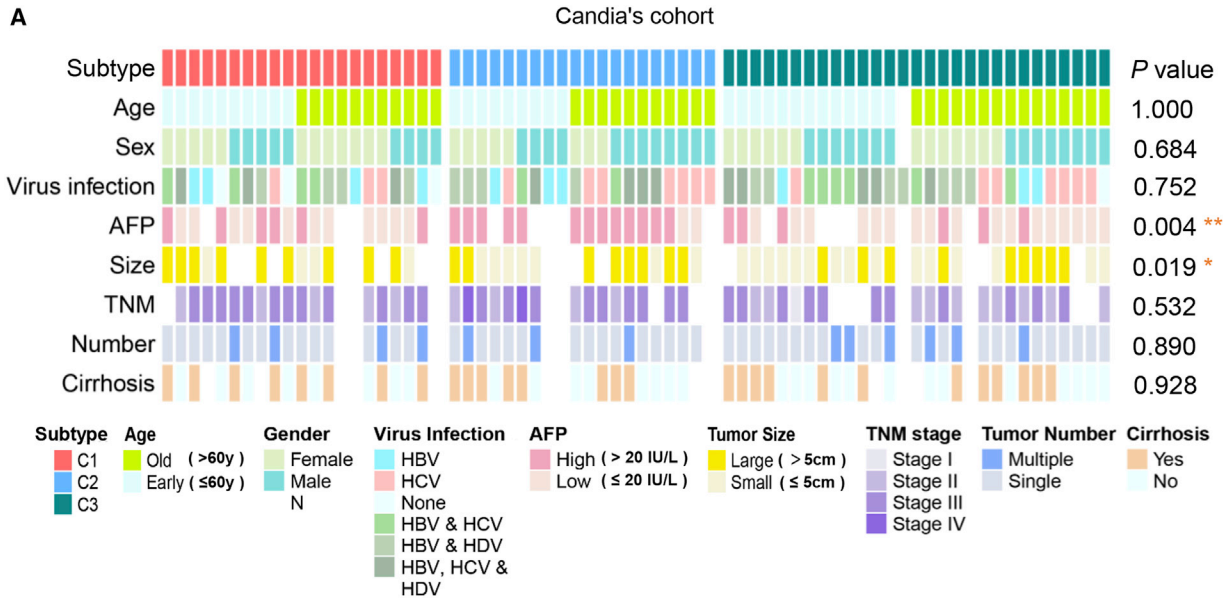
Previous research has identified several HCC subgroups associated with recurrence or OS.[18,19,24–26] In our study, C1 showed great enrichment of metastasis-related pathways, which is associated with the Roessler et al.[19] high-risk subgroup. Notch and transforming growth factor β (TGF-β) signaling were aberrantly activated in C1, which indicates a strong correlation with the Hoshida et al. S1 subgroup. C2 was distinguished by high proliferation and poor prognosis, which is similar to Boyault et al.[18] subtype G2. Individuals in C2 had the highest frequency of TP53 mutation. The TP53 mutation is strongly associated with cell proliferation and chromosome instability, which may partially result in the malignant features of C2. The high AFP level of C2 is in concordance with the "progenitor"-like feature of the Hoshida et al. S2 subgroup.[39] However, a controversial result was observed in the correlation analyses between the Lee et al. subtype and eRNA subtype in different HCC cohorts, which may result from different genetic background and etiology. C3 is strongly linked to the Hoshida et al. S3 subgroup. The good prognosis of C3 is also concordant with Lee et al. subtype B and the Roessler et al.[19] low-risk subgroup. eRNA-based subtyping showed high consistency with published HCC subgroups, which implied that eRNAs were involved in HCC heterogeneity. Dissecting tumor heterogeneity from the perspective of eRNA expression is effective.
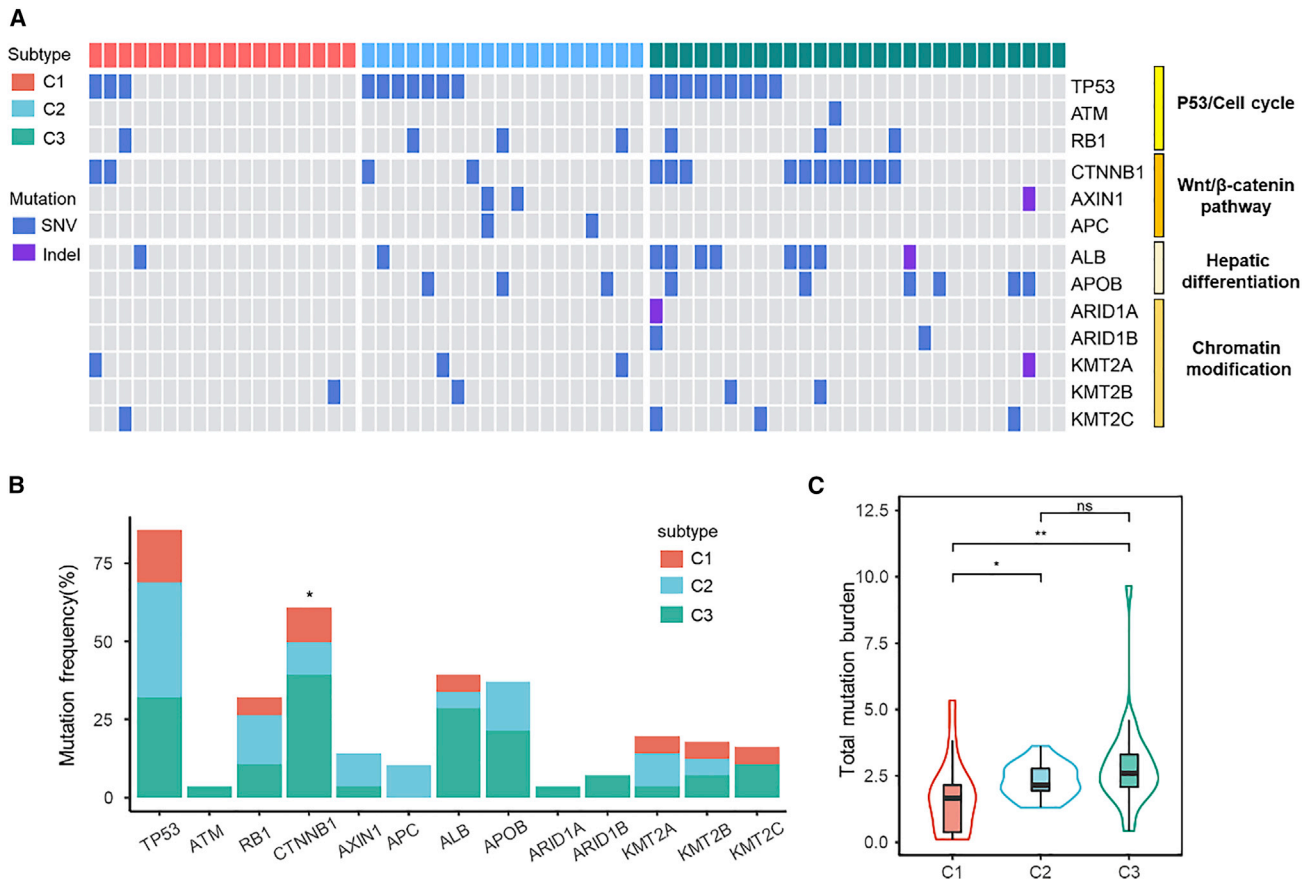
Development of HCC is in part due to multiple complex mutation processes. Aberrant Wnt/β-catenin signaling pathway activation because of CTNNB1 mutation has been observed in more than 20% of individuals with HCC.[40,41] The highest mutant frequency of CTNNB1 was observed in the C3 subgroup (39%) compared with C1 (11.1%) and C2 (10%), suggesting dysregulation of the Wnt/β-catenin pathway in C3. It has been reported that activation of β-catenin signaling can lead to reduced immune infiltration and resistance to ICIs,[42] which is consistent with our results showing that individuals in C3 may not benefit from immunotherapy. However, Wnt signaling pathway-targeted inhibitors may be effective for individuals in C3. A large proportion of C3 individuals carried ALB and APOB mutations, implying a highly differentiated status of C3, which may partially account for the good prognosis.[43] Generally, a low tumor mutation burden (TMB) is associated with a poor response to ICIs.[44–46] However, although the TMB of C1 was lower than those of the other two subtypes, C1 tended to respond to ICIs. C2 and C3 turned out to be insensitive to ICIs, partially because chronic antigen exposure because of high TMB may lead to dysfunction and exhaustion of T cells in C2 and C3.[47] There is no significant difference in neoantigens among three subtypes in the TCGA cohort, suggesting that TMB is not the best predictor of immune infiltration and immunotherapy efficacy in HCC.[48] Reduced immune-mediated cytotoxic and proinflammatory activity in high broad CNA-burden tumors had been observed before and was further proven in the C2 subtype.[49] This finding was likely due to alterations in antigen-presenting machinery and decreased rates of neoantigenic mutations.[36] High CNA burden was also linked to high proliferation, TP53 dysfunction, and DNA repair.[36]

**Figure 6. Correlation of the HCC subtypes with immune infiltration**

(A and B) The immune scores and stromal scores of HCC subtypes in the Candia et al.[16] and Yoon et al.[17] cohorts. (C) The cytolytic activity scores of HCC subtypes in the Candia et al.[16] and Yoon et al.[17] cohorts. (D) Expression level of 15 immune checkpoint genes in three HCC subtypes in two cohorts. (E) Enrichment of immune and stromal cells in three HCC subtypes in two cohorts, analyzed by TIMER, MCP-counter, and xCell, respectively. Boxplot "boxes" indicate the first, second, and third quartiles of the data. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

**A** Candia's cohort

Subtype / Age / Sex / Virus infection / AFP / Size / TNM / Number / Cirrhosis

*P* value
1.000
0.684
0.752
0.004 **
0.019 *
0.532
0.890
0.928

Subtype: C1, C2, C3
Age: Old (>60y), Early (≤60y)
Gender: Female, Male, N
Virus Infection: HBV, HCV, None, HBV & HCV, HBV & HDV, HBV, HCV & HDV
AFP: High (> 20 IU/L), Low (≤ 20 IU/L)
Tumor Size: Large (> 5cm), Small (≤ 5cm)
TNM stage: Stage I, Stage II, Stage III, Stage IV
Tumor Number: Multiple, Single
Cirrhosis: Yes, No

**B** Yoon's cohort

Subtype / Age / Sex / Virus infection / AFP / Size / TNM / Child Pugh / Relapse

*P* value
0.654
0. 445
0.178
0.759
0.021 *
0.573
0.180
0.002 **

Subtype: C1, C2, C3
Age: Old (>60y), Early (≤60y)
Gender: Female, Male
Virus Infection: HBV, HCV, Neither
AFP: High (>200ng/ml), Low (≤ 200ng/ml)
Tumor Size: Large (> 5cm), Small (≤5cm)
TNM stage: Stage I, Stage II, Stage III, Stage IV
Child Pugh: A, B, C
Relapse: Yes, No

**C** Candia's cohort

**D** Yoon's cohort

eRNA subtype: C1, C2, C3
Boyault subtype ***: G1, G2, G3, G4, G5, G6
Hoshida subtype ***: S1, S2, S3
Lee subtype ***: A, B
Roessler subtype ***: High-risk, Low-risk

eRNA subtype: C1, C2, C3
Boyault subtype *: G1, G2, G3, G4, G5, G6
Hoshida subtype ***: S1, S2, S3
Lee subtype **: A, B
Roessler subtype ***: High-risk, Low-risk

*(legend on next page)*

**Figure 8. Genomic features of three HCC subtypes**

(A) Gene mutation landscape of HCC subtypes in the Candia et al.[16] cohort. (B) Mutation frequency of genes in the Candia et al.[16] cohort. (C) Total TMB of HCC subtypes in the Candia et al.[16] cohort. Boxplot "boxes" indicate the first, second, and third quartiles of the data.*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

Targeted therapy and immune checkpoint blockade therapy epitomize new precision treatments for HCC. Sorafenib is a multitargeted kinase inhibitor and currently recommended as the first-line therapy of advanced HCC.[50–53] Another inhibitor of multiple receptor tyrosine kinases, cabozantinib, is a treatment option for individuals with HCC previously treated with sorafenib.[54] As for immunotherapy, PD-1 inhibitors and CTLA-4 inhibitors also have promising applications in HCC.[33,35,55] However, the low response rate and lack of efficacy predictor of ICIs still remain huge challenges. Here we integrated subtype-specific eRNAs into the Candia et al.[16] and Yoon et al.[17] cohorts to obtain a 51-eRNA classifier with the aim of classifying HCC subgroups and predicting responses to different therapies. The validation result showed that the C1, C2, and C3 subtypes were sensitive to ICIs, targeted therapy, and TACE, respectively. The high concordance of classifiers among different cohorts indicated that the eRNA classifier could reproducibly determine HCC classification and guide individualized therapy.

Our study has the following strengths. First, our study provides a new approach for HCC classification based on the expression profile of eRNAs. We utilized total RNA sequencing data to reveal the eRNA landscape in HCC, which is more credible than poly(A) enrichment sequencing. Most importantly, we established and validated a 51-eRNA classifier to predicting prognosis and treatment responses. Nevertheless, the study has some limitations. First, the sample size was relatively small because of the limited numbers of available total RNA sequencing cohorts. Second, comprehensive analyses together with other omics (e.g., proteomics and metabolomics) are needed in the future. Third, the biological roles of eRNAs in HCC carcinogenesis need further exploration.

Our results provide a novel method for immune classification of HCC, shed new light on tumor heterogeneity, and may aid in HCC immunotherapy.

**Figure 7. Clinical characteristics of HCC subtypes in two cohorts**

(A and B) Clinical characteristics of HCC subtypes in the Candia et al.[16] and Yoon et al.[17] cohorts. (C and D) Association of eRNA-based subtypes with previous HCC classifications in the Candia et al.[16] and Yoon et al.[17] cohorts. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
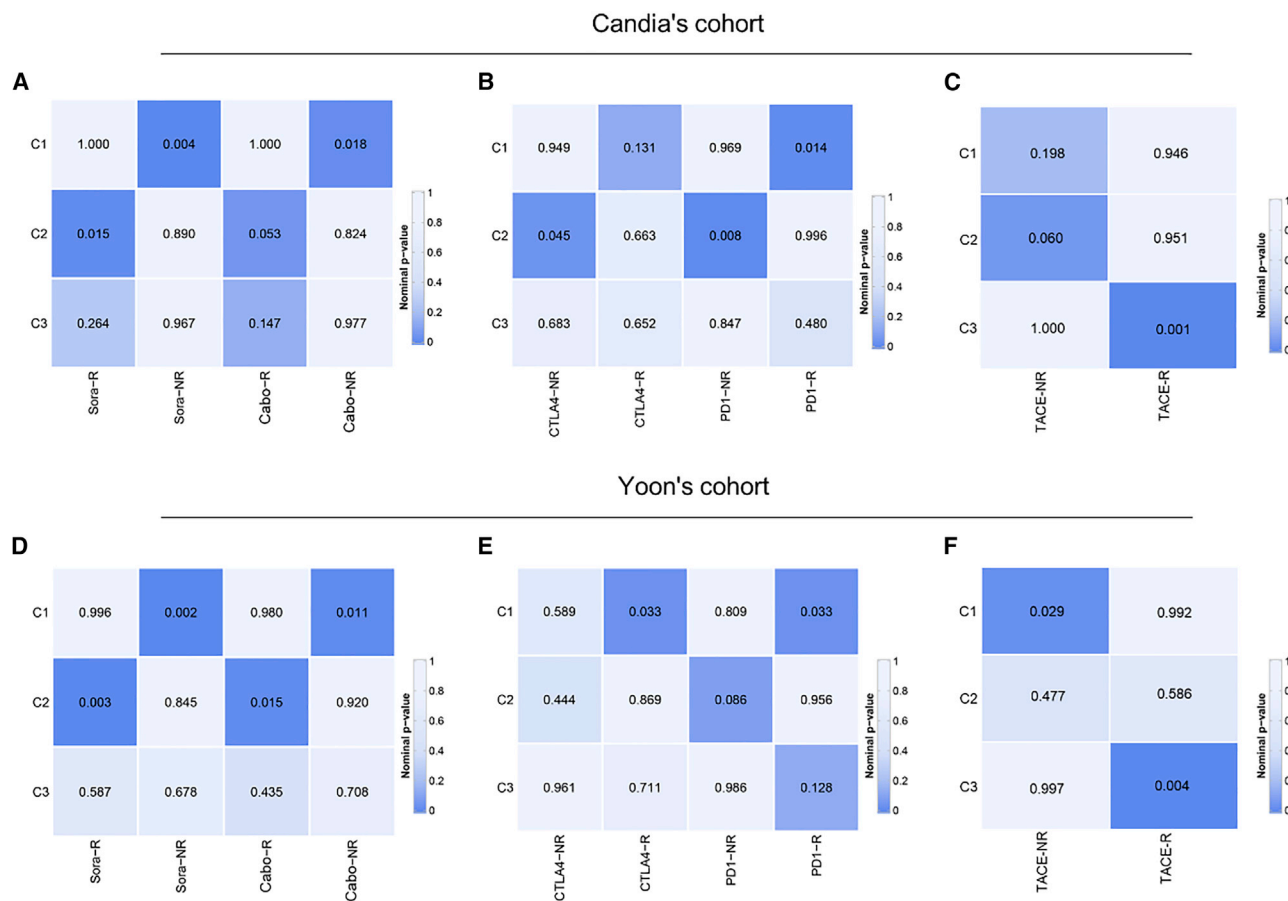
**Figure 9. Correlation of the HCC subtypes with the response to targeted therapy, immunotherapy, and TACE**

(A–C) SubMap correlation analysis between HCC subtypes in the Candia et al.[16] cohort and samples with different sensitivities to targeted therapy (A), ICIs (B), and TACE (C). (D–F) SubMap correlation analysis between HCC subtypes in the Yoon et al.[17] cohort and samples with different sensitivities to targeted therapy (D), ICIs (E), and TACE (F).

## Materials and methods

### RNA sequencing of HCC samples

The institutional review board of SYSUCC approved this study, and written informed consent was obtained from all individuals. Baseline information of individuals in the SYSUCC cohort is shown in Table S1. A total of 14 fresh HCC tumor tissue samples were collected immediately after liver biopsy and stored in RNA stabilization solution (Invitrogen, USA) at −80°C. For preparing the transcriptome library, total RNA was extracted and treated with a rRNA removal kit (Invitrogen). After RNA integrity detection, cDNA synthesis, end repair, A-base addition, and ligation of index adapters were performed. Paired-end libraries were sequenced using the HiSeq X Ten platform.

### Data collection and RNA sequencing analysis

Two groups of total RNA sequencing (RNA-seq) data from the Yoon et al.[17] (GSE148355) and Candia et al.[16] cohorts (GSE144269) were included in our study. Baseline information of individuals in the two cohorts is shown in Table S1. Raw reads in fastq format were fetched from the European Nucleotide Archive (ENA) database using Aspera software. Clinical data of individuals with HCC were obtained from two

previous studies by Yoon et al.[17] and Candia et al.[16] For preprocessing, the sequencing adapters were first removed, and the sequencing quality of the reads was checked using fastp.[56] Then the filtered RNA-seq reads were aligned to the human reference genome hg38 using STAR.[57] Raw gene counts were quantified using featureCounts from the Subread toolkit and then transformed into transcripts per kilobase of exon model per million mapped reads (TPM) using an in-house R script.[58,59]

GEP and somatic mutation data of the TCGA-LIHC cohort were obtained using the R TCGAbiolinks and TCGAmutations package, respectively.[60,61] Somatic copy number variation data and clinical information for the TCGA-LIHC cohort were downloaded from the Firebrowse database (http://firebrowse.org/) and University of California, Santa Cruz (UCSC) Xeno (https://xena.ucsc.edu/public).

For drug sensitivity analyses, the raw microarray data (GSE104580) of individuals with HCC who received TACE therapy were analyzed using the R affy package.[62] Additionally, 34 HCC cell lines with gene expression as well as drug sensitivity data (AUC values) were acquired from the Liver Cancer Cell Lines Database (http://lccl.zucmanlab.com).[63]

*(legend on next page)*

## TCGA



## SYSUCC
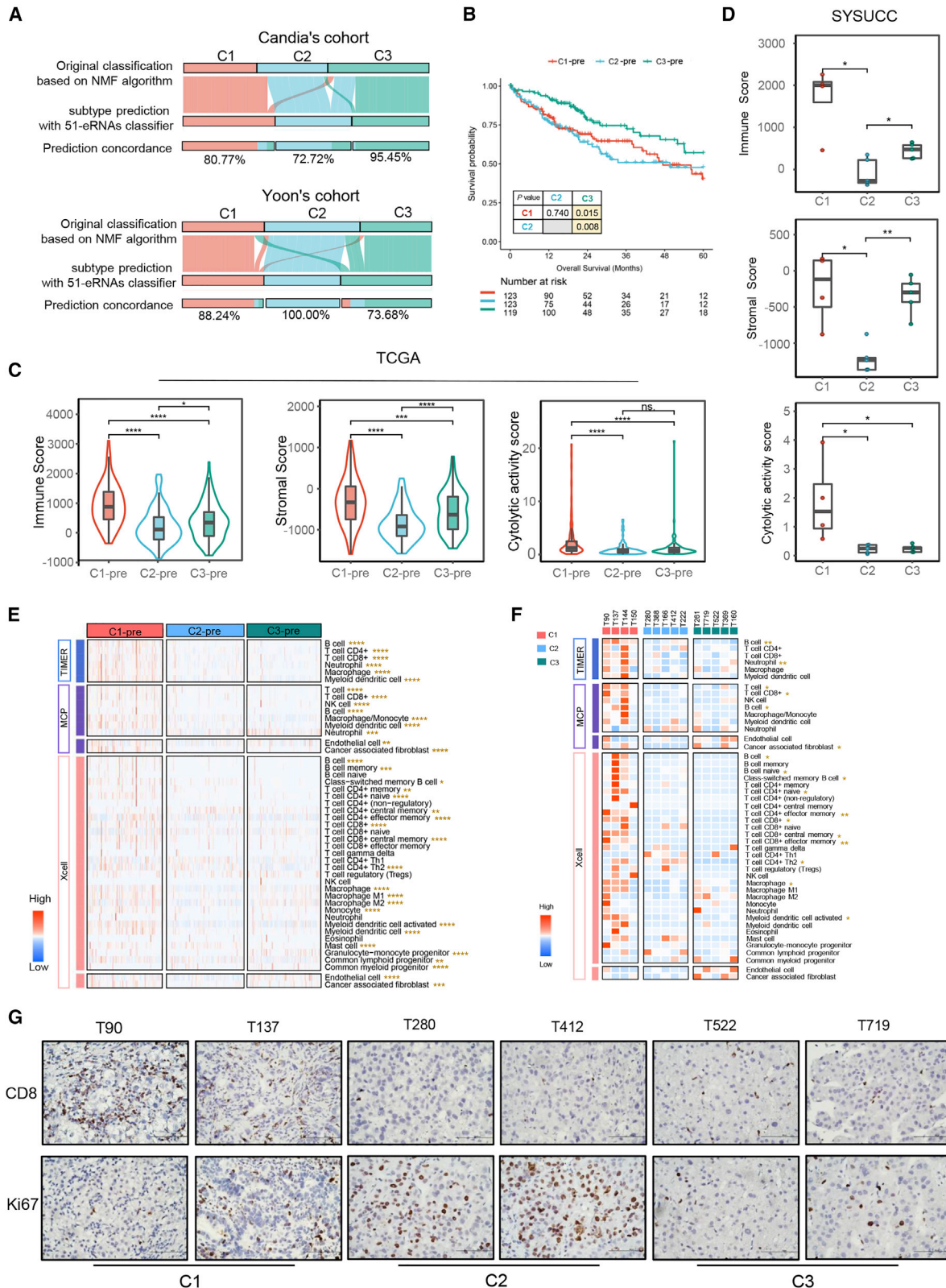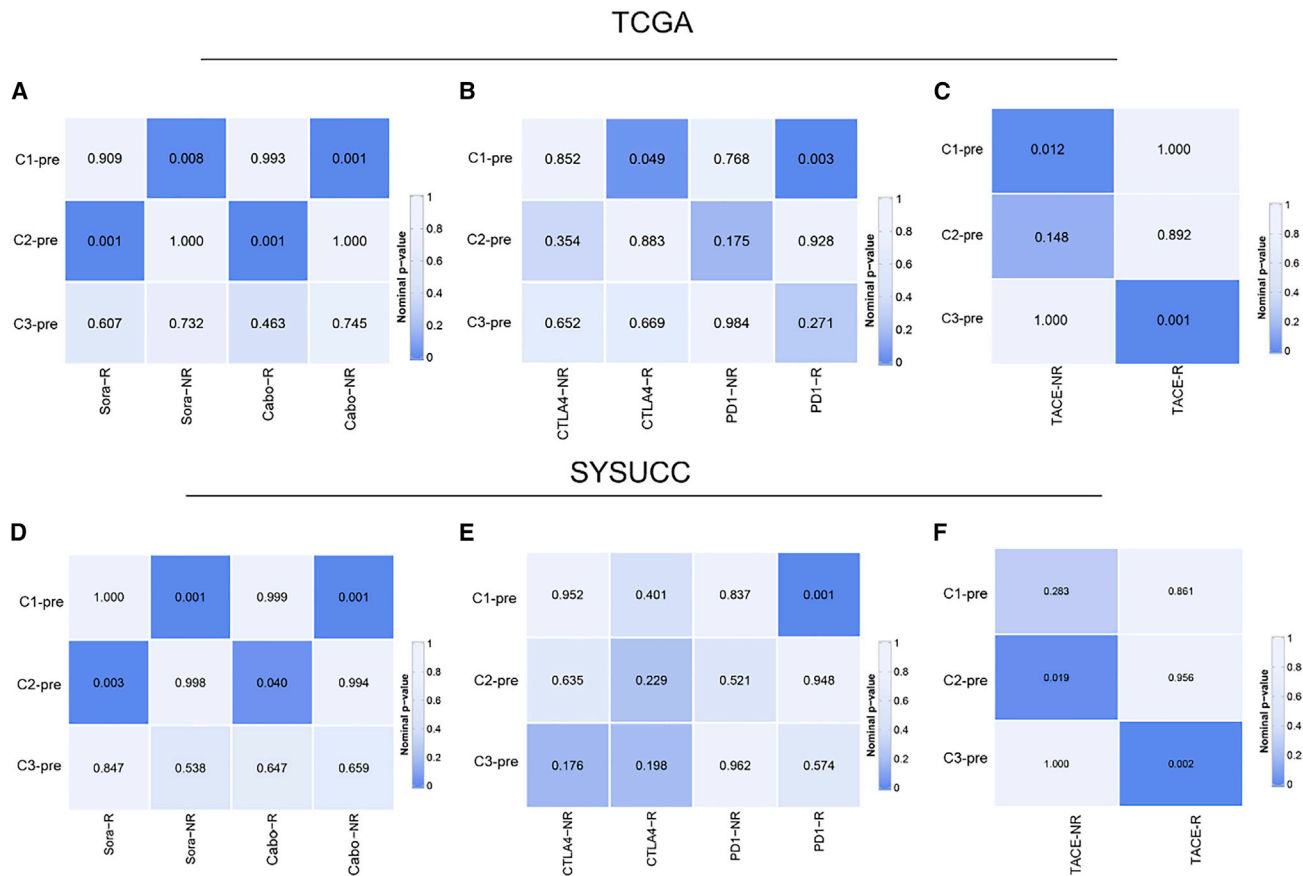


**Figure 11. Predictive value of the 51-eRNA classifier in response to targeted therapy, immunotherapy, and TACE**

(A–C) SubMap correlation analysis between HCC subtypes in the TCGA-LIHC cohort and samples with different sensitivities to targeted therapy (A), ICIs (B), and TACE (C). (D–F) SubMap correlation analysis between HCC subtypes in the SYSUCC cohort and samples with different sensitivities to targeted therapy (D), ICIs (E), and TACE (F).

### eRNA quantification

A previously published pipeline, Pipeline for Enhancer Transcription (PET) (http://fun-science.club/PET/), was adopted for eRNA identification and quantification in our study.[64] Briefly, the trimmed fastq reads were aligned to the hg19 reference genome using STAR, followed by quantification of eRNA using featureCounts with an eRNA annotation file from PET. We calculated reads per kilobase of transcript per million mapped reads (RPKM) values from raw count data and filtered out eRNAs whose length was less than 50 bp to avoid mismatches and multi-mapped transcripts, as reported previously.[64] The detected eRNAs were defined as RPKM > 0 in more than 10% of tumor or normal samples and subjected to analyses. The

coordinate distribution of eRNAs was illustrated by the R circlize package.[65]

### eRNA differential expression analysis and functional annotation

As reported in a previous study, we estimated the difference of eRNA expression at the single and global level.[15] Global eRNA expression was evaluated by aggregating the expression level of single eRNAs in each sample and then averaged by the number of detected eRNAs. For single-eRNA-level analysis, the difference in $\log_2$-transformed RPKM value of each eRNA between two groups was evaluated by Wilcoxon test, and the p value was corrected by false discovery rate (FDR) using the

**Figure 10. Construction and validation of the subtype-specific eRNA classifier**

(A) Concordance of HCC subtypes between the prediction based on the 51-eRNA classifier and original classification based on NMF in two cohorts. (B) OS of three HCC subtypes predicted by the 51-eRNA classifier in the TCGA-LIHC cohort.

(C and D) The immune scores, stromal scores, and cytolytic activity scores of HCC subtypes predicted by the 51-eRNA classifier in the TCGA-LIHC cohort (C) and SYSUCC cohort (D). (E and F) Enrichment of immune and stromal cells in three HCC subtypes predicted by the 51-eRNA classifier in the TCGA-LIHC cohort (E) and SYSUCC cohort (F), analyzed by TIMER, MCP, and Xcell, respectively. (G) Representative images of IHC-stained slides with CD8 and ki67 antibodies in three subtypes in the SYSUCC cohort. Scale bars, 100 μm. Boxplot "boxes" indicate the first, second, and third quartiles of the data. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

**Figure 12. Overview of three HCC subtypes based on eRNA expression profile**

Benjamini-Hochberg method for multiple-sample tests.[66] A single eRNA with an absolute $\log_2$FC (fold change) value greater than 1 and FDR less than 0.05 was defined as differentially expressed.

To annotate the biological functions of the upregulated eRNAs ($\log_2$FC > 1 and FDR < 0.05) in tumors, we first performed Spearman co-expression analysis between eRNAs and PCGs. PCGs with a correlation coefficient greater than 0.6 and p value less than 0.05 were considered positively co-expressed. Then we estimated the enriched KEGG pathways (adjusted p < 0.05) of the co-expressed genes with the R clusterProfiler package.[67]

### Identification of eRNA-associated HCC subtypes

NMF subtyping was conducted to subtype the eRNA-related HCC subclasses in both cohorts by the R NMF package.[68] We first screened out eRNAs that were detected in over 50% of individuals and then chose the first 50% of eRNAs (n = 1,244) with a higher median absolute deviation value in the Candia et al.[16] cohort to perform the NMF analysis. The subtyping procedure was also applied to the Yoon et al.[17] cohort with most of the same candidate eRNAs as in the Candia et al.[16] cohort, except for 9 eRNAs that were not detected in the Yoon et al.[17] cohort. The eRNAs used in the NMF analysis are shown in Table S3. In the NMF subtyping procedure with 30 iterations, the optimal number of subtypes was determined by determining the cophenetic score (reproducibility of a model) from 2 to 7 subtypes where the value of the cophenetic score started to fall sharply.[69] PCA was conducted to validate the subtype assignments. Based on the eRNA expression profile, subclass mapping analysis (Submap) was carried out to evaluate the similarity of eRNA subtypes from the two cohorts.[70]

### Molecular characterization and immune infiltration estimation of HCC subtypes

Single-sample gene set enrichment analysis (ssGSEA), based on 50 cancer hallmark gene sets and 186 KEGG gene sets, was performed to evaluate the biological properties of HCC subtypes with the R GSVA package.[71] Previously published HCC molecular subtypes were predicted by the MS.liverK algorithm.[72]

To illustrate the immune infiltration of HCC subtypes, we first compared the immune score and stromal score among different HCC subtypes by applying the R estimate package.[73] Then the TME landscape was characterized by inferring the abundance of immune and stromal (endothelial and fibroblast) cells from the transcriptome data via the TIMER2 online tool.[74–76] We calculated the cytolytic activity score using the MCP-counter algorithm.[75]

### Prediction of the response to immunotherapy, targeted therapy, and TACE for HCC subtypes

Based on SubMap analysis, we first evaluated the similarity of gene expression patterns between HCC subtypes and previously reported individuals with melanoma[77] who received CTLA-4 blockade or PD-1 blockade immunotherapy to infer the probability of benefits from ICIs. Similarly, the sensitivity to two targeted drugs, sorafenib and cabozantinib, was evaluated according to the data of HCC cell lines.[63] To be more specific, the HCC cell lines were arranged from low to high according to the drug AUC value, and cell lines in the first one-third were regarded as drug sensitive, whereas the last one-third were drug resistant. TACE therapeutic efficacy was also estimated for the identified subtypes based on our microarray data.

### Identification of eRNA signature for HCC subtype prediction

For HCC subtype prediction, we first performed differential expression analyses between one subtype and the other two subtypes in every cohort to identify the subtype-specific eRNAs. Only eRNAs with a significant difference ($\log_2 FC > 1$, FDR < 0.05) were defined as subtype-specific eRNAs. We also identified subtype-specific eRNAs that were shared by two associated subtypes from the two cohorts as the classifier eRNAs. As a result, a 51-eRNA classifier was established. To test the reliability of the classifier, the concordance of predictive subtype, through nearest template prediction (NTP) analysis,[78] was compared with previously identified subtypes based on NMF in the Yoon et al.[17] and Candia et al.[16] cohorts, respectively.

### Validation of the eRNA subtypes in the TCGA and SYSUCC HCC cohorts

Using the NTP algorithm, we classified individuals with HCC from the TCGA-LIHC and SYSUCC HCC cohorts into three subtypes based on the 51-eRNA classifier. The eRNA expression profile of the TCGA-LIHC cohort was downloaded from the Cancer eRNA Atlas (https://bioinformatics.mdanderson.org/public-software/tcea).[79] Survival difference, molecular characterization, immune infiltration, and drug sensitivity analyses were also performed as described above.

### Correlation of eRNA subtypes with clinical features, CNA burden, gene mutations, and neoantigens

Correlations between clinical features and HCC subtypes were analyzed in the Candia et al.[16] and Yoon et al.[17] cohorts. For gene somatic mutation data analysis in the Candia et al.[16] and TCGA-LIHC cohorts, TMB was calculated with the R Maftools package. The mutation frequency of several HCC-related genes was compared.[80] The neoantigens data for individuals with HCC in the TCGA-LIHC cohort were acquired from a previous study.[81] CNA burden was quantified by CNApp (https://tools.idibaps.org/CNApp/).[82]

### IHC

HCC samples were fixed in 4% formaldehyde for 24 h at room temperature, moved into 70% ethanol for 12 h, and then embedded in paraffin. After cutting and baking at 60°C for 60 min for de-paraffinization, the samples were incubated in graded alcohol and then treated for antigen unmasking. For IHC staining, endogenous peroxidases were inactivated by 3% hydrogen peroxide at room temperature (RT) for 15 min. The slides were treated with Tris-EDTA buffer (10 mM, pH 8.0) at 121°C for 10 min for antigen retrieval. Tissues were stained with primary antibodies for 12 h at 4°C. After washing with PBS-Tween, tissues were stained with horseradish peroxidase (HRP)-conjugated secondary antibodies against mouse or rabbit for 30 min at 37°C. Then the samples were co-stained with hematoxylin, followed by washing with PBS. The dilution ratios for IHC were 1:500 for CD8 antibodies (ZSGB-BIO, ZA-0508), 1:500 for ki67 antibodies (Invitrogen, PA5-114437), and 1:500 for HRP-conjugated secondary antibodies (ZSGB-BIO, DS-0003).

### Statistical analysis

All statistical analyses and data illustrations were conducted using R software (v.4.0.0). Kaplan-Meier survival analysis and log rank test between different subtypes were performed by utilizing the pairwise_survdiff() function in the R survival package. Cox proportional hazards modeling and ROC curve analyses were conducted using the R survminer and timeROC packages, respectively. Chi-square test or Fisher's exact test was used to test the difference in categorical data. For differences in continuous data between two or multiple groups, Wilcoxon test or Kruskal-Wallis test was performed. The p values were corrected by FDR using the Benjamini-Hochberg method when multiple comparisons occurred. A two-tailed p value od less than 0.05 was regarded as statistically significant.

## AVAILABILITY OF DATA

The TCGA-LIHC and GEO datasets used in this study are available in public databases. The data of the SYSUCC cohort is available from the corresponding authors upon reasonable request.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omto.2022.07.001.

## ACKNOWLEDGMENTS

## AUTHOR CONTRIBUTIONS

Conception and design, X.B., M.S., and S.L.; collection and assembly of data, X.B. and S.L.; data analysis and interpretation, X.B. and S.L.; manuscript writing, X.B., X.L., and Y.X.; manuscript editing, A.K., D.W., and L.S.; final approval of manuscript, all authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA. Cancer J. Clin. 65, 87–108. https://doi.org/10.3322/caac.21262.

2. Bruix, J., Gores, G.J., and Mazzaferro, V. (2014). Hepatocellular carcinoma: clinical frontiers and perspectives. Gut 63, 844–855. https://doi.org/10.1136/gutjnl-2013-306627.

3. Craig, A.J., von Felden, J., Garcia-Lezana, T., Sarcognato, S., and Villanueva, A. (2020). Tumour evolution in hepatocellular carcinoma. Nat. Rev. Gastroenterol. Hepatol. 17, 139–152. https://doi.org/10.1038/s41575-019-0229-4.

4. de Bono, J.S., and Ashworth, A. (2010). Translating cancer research into targeted therapeutics. Nature 467, 543–549. https://doi.org/10.1038/nature09339.

5. Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. 15, 272–286. https://doi.org/10.1038/nrg3682.

6. Ong, C.T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293. https://doi.org/10.1038/nrg2957.

7. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74. https://doi.org/10.1038/nature11247.

8. Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet. *20*, 437–455. https://doi.org/10.1038/s41576-019-0128-0.

9. Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327–339. https://doi.org/10.1016/j.cell.2011.01.024.

10. Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390–394. https://doi.org/10.1038/nature10006.

11. Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet. *17*, 207–223. https://doi.org/10.1038/nrg.2016.4.

12. Zhang, Z., Lee, J.H., Ruan, H., Ye, Y., Krakowiak, J., Hu, Q., Xiang, Y., Gong, J., Zhou, B., Wang, L., et al. (2019). Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. Nat. Commun. *10*, 4562. https://doi.org/10.1038/s41467-019-12543-5.

13. Tan, S.H., Leong, W.Z., Ngoc, P.C.T., Tan, T.K., Bertulfo, F.C., Lim, M.C., An, O., Li, Z., Yeoh, A.E.J., Fullwood, M.J., et al. (2019). The enhancer RNA ARIEL activates the oncogenic transcriptional program in T-cell acute lymphoblastic leukemia. Blood *134*, 239–251. https://doi.org/10.1182/blood.2018874503.

14. Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520. https://doi.org/10.1038/nature12210.

15. Qin, N., Ma, Z., Wang, C., Zhang, E., Li, Y., Huang, M., Chen, C., Zhang, C., Fan, J., Gu, Y., et al. (2020). Comprehensive characterization of functional eRNAs in lung adenocarcinoma reveals novel regulators and a prognosis-related molecular subtype. Theranostics *10*, 11264–11277. https://doi.org/10.7150/thno.47039.

16. Candia, J., Bayarsaikhan, E., Tandon, M., Budhu, A., Forgues, M., Tovuu, L.O., Tudev, U., Lack, J., Chao, A., Chinburen, J., and Wang, X.W. (2020). The genomic landscape of Mongolian hepatocellular carcinoma. Nat. Commun. *11*, 4383. https://doi.org/10.1038/s41467-020-18186-1.

17. Yoon, S.H., Choi, S.W., Nam, S.W., Lee, K.B., and Nam, J.W. (2021). Preoperative immune landscape predisposes adverse outcomes in hepatocellular carcinoma patients with liver transplantation. NPJ Precis. Oncol. *5*, 27. https://doi.org/10.1038/s41698-021-00167-2.

18. Boyault, S., Rickman, D.S., de Reyniès, A., Balabaud, C., Rebouissou, S., Jeannot, E., Hérault, A., Saric, J., Belghiti, J., Franco, D., et al. (2007). Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology *45*, 42–52. https://doi.org/10.1002/hep.21467.

19. Roessler, S., Jia, H.L., Budhu, A., Forgues, M., Ye, Q.H., Lee, J.S., Thorgeirsson, S.S., Sun, Z., Tang, Z.Y., Qin, L.X., and Wang, X.W. (2010). A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer. Res. *70*, 10202–10212. https://doi.org/10.1158/0008-5472.CAN-10-2607.

20. Llovet, J.M. (2005). Updated treatment approach to hepatocellular carcinoma. J. Gastroenterol. *40*, 225–235. https://doi.org/10.1007/s00535-005-1566-3.

21. Verslype, C., Rosmorduc, O., and Rougier, P.; ESMO Guidelines Working Group (2012). Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol. *23*, vii41–vii48. https://doi.org/10.1093/annonc/mds225.

22. Edge, S.B., and Compton, C.C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann. Surg. Oncol. *17*, 1471–1474. https://doi.org/10.1245/s10434-010-0985-4.

23. Llovet, J.M., Brú, C., and Bruix, J. (1999). Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin. Liver. Dis. *19*, 329–338. https://doi.org/10.1055/s-2007-1007122.

24. Lee, J.S., Chu, I.S., Heo, J., Calvisi, D.F., Sun, Z., Roskams, T., Durnez, A., Demetris, A.J., and Thorgeirsson, S.S. (2004). Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology *40*, 667–676. https://doi.org/10.1002/hep.20375.

25. Lee, J.S., and Thorgeirsson, S.S. (2004). Genome-scale profiling of gene expression in hepatocellular carcinoma: classification, survival prediction, and identification of therapeutic targets. Gastroenterology *127*, S51–S55. https://doi.org/10.1053/j.gastro.2004.09.015.

26. Hoshida, Y., Nijman, S.M.B., Kobayashi, M., Chan, J.A., Brunet, J.P., Chiang, D.Y., Villanueva, A., Newell, P., Ikeda, K., Hashimoto, M., et al. (2009). Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer. Res. *69*, 7385–7392. https://doi.org/10.1158/0008-5472.CAN-09-1089.

27. Lam, M.T.Y., Li, W., Rosenfeld, M.G., and Glass, C.K. (2014). Enhancer RNAs and regulated transcriptional programs. Trends. Biochem. Sci. *39*, 170–182. https://doi.org/10.1016/j.tibs.2014.02.007.

28. Hou, T.Y., and Kraus, W.L. (2021). Spirits in the material world: enhancer RNAs in transcriptional regulation. Trends. Biochem. Sci. *46*, 138–153. https://doi.org/10.1016/j.tibs.2020.08.007.

29. Lee, J.H., Xiong, F., and Li, W. (2020). Enhancer RNAs in cancer: regulation, mechanisms and therapeutic potential. RNA Biol. *17*, 1550–1559. https://doi.org/10.1080/15476286.2020.1712895.

30. Fu, Y., Liu, S., Zeng, S., and Shen, H. (2019). From bench to bed: the tumor immune microenvironment and current immunotherapeutic strategies for hepatocellular carcinoma. J. Exp. Clin. Cancer. Res. *38*, 396. https://doi.org/10.1186/s13046-019-1396-4.

31. Kurebayashi, Y., Ojima, H., Tsujikawa, H., Kubota, N., Maehara, J., Abe, Y., Kitago, M., Shinoda, M., Kitagawa, Y., and Sakamoto, M. (2018). Landscape of immune microenvironment in hepatocellular carcinoma and its additional impact on histological and molecular classification. Hepatology *68*, 1025–1041. https://doi.org/10.1002/hep.29904.

32. Duffy, A.G., Ulahannan, S.V., Makorova-Rusher, O., Rahma, O., Wedemeyer, H., Pratt, D., Davis, J.L., Hughes, M.S., Heller, T., ElGindi, M., et al. (2017). Tremelimumab in combination with ablation in patients with advanced hepatocellular carcinoma. J. Hepatol. *66*, 545–551. https://doi.org/10.1016/j.jhep.2016.10.029.

33. El-Khoueiry, A.B., Sangro, B., Yau, T., Crocenzi, T.S., Kudo, M., Hsu, C., Kim, T.Y., Choo, S.P., Trojan, J., Welling, T.H., Rd, et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet *389*, 2492–2502. https://doi.org/10.1016/S0140-6736(17)31046-2.

34. Killock, D. (2017). Immunotherapy: nivolumab keeps HCC in check and opens avenues for checkmate. Nat. Rev. Clin. Oncol. *14*, 392. https://doi.org/10.1038/nrclinonc.2017.70.

35. Zhu, A.X., Finn, R.S., Edeline, J., Cattan, S., Ogasawara, S., Palmer, D., Verslype, C., Zagonel, V., Fartoux, L., Vogel, A., et al.; KEYNOTE-224 investigators (2018). Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet. Oncol. *19*, 940–952. https://doi.org/10.1016/S1470-2045(18)30351-6.

36. Bassaganyas, L., Pinyol, R., Esteban-Fabró, R., Torrens, L., Torrecilla, S., Willoughby, C.E., Franch-Expósito, S., Vila-Casadesús, M., Salaverria, I., Montal, R., et al. (2020). Copy-number alteration burden differentially impacts immune profiles and molecular features of hepatocellular carcinoma. Clin. Cancer. Res. *26*, 6350–6361. https://doi.org/10.1158/1078-0432.CCR-20-1497.

37. Chen, Y., Li, Z.Y., Zhou, G.Q., and Sun, Y. (2021). An immune-related gene prognostic index for head and neck squamous cell carcinoma. Clin. Cancer. Res. *27*, 330–341. https://doi.org/10.1158/1078-0432.CCR-20-2166.

38. Pinyol, R., Montal, R., Bassaganyas, L., Sia, D., Takayama, T., Chau, G.Y., Mazzaferro, V., Roayaie, S., Lee, H.C., Kokudo, N., et al. (2019). Molecular predictors of prevention of recurrence in HCC with sorafenib as adjuvant treatment and prognostic factors in the phase 3 STORM trial. Gut *68*, 1065–1075. https://doi.org/10.1136/gutjnl-2018-316408.

39. Laurent-Puig, P., Legoix, P., Bluteau, O., Belghiti, J., Franco, D., Binot, F., Monges, G., Thomas, G., Bioulac-Sage, P., and Zucman-Rossi, J. (2001). Genetic alterations

associated with hepatocellular carcinomas define distinct pathways of hepatocarcino-genesis. Gastroenterology *120*, 1763–1773. https://doi.org/10.1053/gast.2001.24798.

40. Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and path-ways in hepatocellular carcinoma. Nat. Genet. *44*, 694–698. https://doi.org/10.1038/ng.2256.

41. Delgado, E., Okabe, H., Preziosi, M., Russell, J.O., Alvarado, T.F., Oertel, M., Nejak-Bowen, K.N., Zhang, Y., and Monga, S.P.S. (2015). Complete response of Ctnnb1-mutated tumours to beta-catenin suppression by locked nucleic acid antisense in a mouse hepatocarcinogenesis model. J. Hepatol. *62*, 380–387. https://doi.org/10.1016/j.jhep.2014.10.021.

42. Spranger, S., Bao, R., and Gajewski, T.F. (2015). Melanoma-intrinsic beta-catenin sig-nalling prevents anti-tumour immunity. Nature *523*, 231–235. https://doi.org/10.1038/nature14404.

43. Cancer Genome Atlas Research Network Electronic address wheeler@bcm.edu. Electronic address, w.b.e.; Cancer Genome Atlas Research, N (2017). Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell *169*, 1327–1341.e23. e1323. https://doi.org/10.1016/j.cell.2017.05.046.

44. Samstein, R.M., Lee, C.H., Shoushtari, A.N., Hellmann, M.D., Shen, R., Janjigian, Y.Y., Barron, D.A., Zehir, A., Jordan, E.J., Omuro, A., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat. Genet. *51*, 202–206. https://doi.org/10.1038/s41588-018-0312-8.

45. Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., et al. (2015). Cancer immunology. Mutational land-scape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science *348*, 124–128. https://doi.org/10.1126/science.aaa1348.

46. Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J.M., Desrichard, A., Walsh, L.A., Postow, M.A., Wong, P., Ho, T.S., et al. (2014). Genetic basis for clinical response to CTLA-4 blockade in melanoma. N. Engl. J. Med. *371*, 2189–2199. https://doi.org/10.1056/NEJMoa1406498.

47. Ghorani, E., Reading, J.L., Henry, J.Y., Massy, M.R.d., Rosenthal, R., Turati, V., Joshi, K., Furness, A.J.S., Ben Aissa, A., Saini, S.K., et al. (2020). The T cell differentiation landscape is shaped by tumour mutations in lung cancer. Nat. Cancer *1*, 546–561. https://doi.org/10.1038/s43018-020-0066-y.

48. Lu, L., Jiang, J., Zhan, M., Zhang, H., Wang, Q.T., Sun, S.N., Guo, X.K., Yin, H., Wei, Y., Liu, J.O., et al. (2021). Targeting neoantigens in hepatocellular carcinoma for immunotherapy: a futile strategy? Hepatology *73*, 414–421. https://doi.org/10.1002/hep.31279.

49. Davoli, T., Uno, H., Wooten, E.C., and Elledge, S.J. (2017). Tumor aneuploidy corre-lates with markers of immune evasion and with reduced response to immunotherapy. Science *355*, eaaf8399. https://doi.org/10.1126/science.aaf8399.

50. Llovet, J.M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.F., de Oliveira, A.C., Santoro, A., Raoul, J.L., Forner, A., et al.; SHARP Investigators Study Group (2008). Sorafenib in advanced hepatocellular carcinoma. N. Engl. J. Med. *359*, 378–390. https://doi.org/10.1056/NEJMoa0708857.

51. Cheng, A.L., Kang, Y.K., Chen, Z., Tsao, C.J., Qin, S., Kim, J.S., Luo, R., Feng, J., Ye, S., Yang, T.S., et al. (2009). Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet. Oncol. *10*, 25–34. https://doi.org/10.1016/S1470-2045(08)70285-7.

52. Bruix, J., Raoul, J.L., Sherman, M., Mazzaferro, V., Bolondi, L., Craxi, A., Galle, P.R., Santoro, A., Beaugrand, M., Sangiovanni, A., et al. (2012). Efficacy and safety of sor-afenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial. J. Hepatol. *57*, 821–829. https://doi.org/10.1016/j.jhep.2012.06.014.

53. Cheng, A.L., Guan, Z., Chen, Z., Tsao, C.J., Qin, S., Kim, J.S., Yang, T.S., Tak, W.Y., Pan, H., Yu, S., et al. (2012). Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma according to baseline status: subset analyses of the phase III Sorafenib Asia-Pacific trial. Eur. J. Cancer. *48*, 1452–1465. https://doi.org/10.1016/j.ejca.2011.12.006.

54. Abou-Alfa, G.K., Meyer, T., Cheng, A.L., El-Khoueiry, A.B., Rimassa, L., Ryoo, B.Y., Cicin, I., Merle, P., Chen, Y., Park, J.W., et al. (2018). Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. N. Engl. J. Med. *379*, 54–63. https://doi.org/10.1056/NEJMoa1717002.

55. Sangro, B., Gomez-Martin, C., de la Mata, M., Iñarrairaegui, M., Garralda, E., Barrera, P., Riezu-Boj, J.I., Larrea, E., Alfaro, C., Sarobe, P., et al. (2013). A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C. J. Hepatol. *59*, 81–88. https://doi.org/10.1016/j.jhep.2013.02.022.

56. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

57. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

58. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

59. Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scal-able read mapping by seed-and-vote. Nucleic. Acids. Res. *41*, e108. https://doi.org/10.1093/nar/gkt214.

60. Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic. Acids. Res. *44*, e71. https://doi.org/10.1093/nar/gkv1507.

61. Ellrott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C., Hess, J., Ma, S., Chiotti, K.E., McLellan, M., et al.; MC3 Working Group; Cancer Genome Atlas Research Network (2018). Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. Cell Syst. *6*, 271–281.e7. e277. https://doi.org/10.1016/j.cels.2018.03.002.

62. Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307–315. https://doi.org/10.1093/bioinformatics/btg405.

63. Caruso, S., Calatayud, A.L., Pilet, J., La Bella, T., Rekik, S., Imbeaud, S., Letouzé, E., Meunier, L., Bayard, Q., Rohr-Udilova, N., et al. (2019). Analysis of liver cancer cell lines identifies agents with likely efficacy against hepatocellular carcinoma and markers of response. Gastroenterology *157*, 760–776. https://doi.org/10.1053/j.gastro.2019.05.001.

64. Wu, Y., Yang, Y., Gu, H., Tao, B., Zhang, E., Wei, J., Wang, Z., Liu, A., Sun, R., Chen, M., et al. (2020). Multi-omics analysis reveals the functional transcription and poten-tial translation of enhancers. Int. J. Cancer. *147*, 2210–2224. https://doi.org/10.1002/ijc.33132.

65. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

66. Benjamini Y, H.Y. (1995). Controlling the false discovery rate: a practical and power-ful approach to multiple testing. J. R. Stat. Soc. *57*, 289–300.

67. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.

68. Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. BMC. Bioinformatics. *11*, 367. https://doi.org/10.1186/1471-2105-11-367.

69. Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and mo-lecular pattern discovery using matrix factorization. Proc. Natl. Acad. Sci. USA *101*, 4164–4169. https://doi.org/10.1073/pnas.0308531101.

70. Hoshida, Y., Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2007). Subclass mapping: identifying common subtypes in independent disease data sets. PLoS One *2*, e1195. https://doi.org/10.1371/journal.pone.0001195.

71. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation anal-ysis for microarray and RNA-seq data. BMC. Bioinformatics. *14*, 7. https://doi.org/10.1186/1471-2105-14-7.

72. Petitprez, F. (2019). MS.liverK: an R package for transcriptome-based computation of molecular subtypes and functional signatures in liver cancer. Preprint at BioRxiv. https://doi.org/10.1101/540005.

73. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612. https://doi.org/10.1038/ncomms3612.

74. Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic. Acids. Res. *48*, W509–W514. https://doi.org/10.1093/nar/gkaa407.

75. Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., and de Reyniès, A. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome. Biol. *17*, 218. https://doi.org/10.1186/s13059-016-1070-5.

76. Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome. Biol. *18*, 220. https://doi.org/10.1186/s13059-017-1349-1.

77. Roh, W., Chen, P.L., Reuben, A., Spencer, C.N., Prieto, P.A., Miller, J.P., Gopalakrishnan, V., Wang, F., Cooper, Z.A., Reddy, S.M., et al. (2017). Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. Sci. Transl. Med. *9*, eaah3560. https://doi.org/10.1126/scitranslmed.aah3560.

78. Hoshida, Y. (2010). Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. PLoS One *5*, e15543. https://doi.org/10.1371/journal.pone.0015543.

79. Chen, H., and Liang, H. (2020). A high-resolution map of human enhancer RNA loci characterizes super-enhancer activities in cancer. Cancer. Cell, *38*, 701–715.e5. e705. https://doi.org/10.1016/j.ccell.2020.08.020.

80. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome. Res. *28*, 1747–1756. https://doi.org/10.1101/gr.239244.118.

81. Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell *160*, 48–61. https://doi.org/10.1016/j.cell.2014.12.033.

82. Franch-Expósito, S., Bassaganyas, L., Vila-Casadesús, M., Hernández-Illán, E., Esteban-Fabró, R., Díaz-Gay, M., Lozano, J.J., Castells, A., Llovet, J.M., Castellví-Bel, S., and Camps, J. (2020). CNApp, a tool for the quantification of copy number alterations and integrative analysis revealing clinical implications. Elife *9*, e50267. https://doi.org/10.7554/eLife.50267.