








ORIGINAL ARTICLE

Molecular and radiopathologic spectrum between HCC and intrahepatic cholangiocarcinoma

Youngsic Jeon^{1,2}  | So Mee Kwon³  | Hyungjin Rhee⁴  | Jeong Eun Yoo¹  | Taek Chung⁵  | Hyun Goo Woo^{3,6}  | Young Nyun Park¹ 

¹Department of Pathology, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, Republic of Korea

²Natural Products Research Center, Korea Institute of Science and Technology, Gangneung, Republic of Korea

³Department of Physiology, Ajou University School of Medicine, Suwon, Republic of Korea

⁴Department of Radiology, Yonsei University College of Medicine, Seoul, Republic of Korea

⁵Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

⁶Department of Biomedical Science, Graduate School, Ajou University, Suwon, Republic of Korea

Correspondence

Hyun Goo Woo, Department of Physiology, Ajou University School of Medicine, 164 Worldcup-ro, Yeongtong-gu, Suwon, 16499, Republic of Korea. Email: hg@ajou.ac.kr

Young Nyun Park, Department of Pathology, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea. Email: young0608@yuhs.ac

Funding information

Supported by grants from the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT, Republic of Korea (NRF-2017R1E1A1A01074733, NRF-2017M3A9B6061509, NRF-2017M3C9A6047620, NRF-2017R1A2B4005871, NRF-2017M3A9B6061512, and NRF-2020R1A2B5B01001646) and the Korea Health Industry Development Institute funded by the Ministry of Health and Welfare, Republic of Korea (HI21C1003). The Korea Research Environment Open NETwork supported the data analysis operations, managed and operated by the

Abstract

Background and Aims: Primary liver cancers (LCs), including HCC and intrahepatic cholangiocarcinoma (iCCA), are derived from a common developmental lineage, conferring a molecular spectrum between them. To elucidate the molecular spectrum, we performed an integrative analysis of transcriptome profiles associated with patients' radiopathologic features.

Approach and Results: We identified four LC subtypes (LC1–LC4) from RNA-sequencing profiles, revealing intermediate subtypes between HCC and iCCA. LC1 is a typical HCC characterized by active bile acid metabolism, telomerase reverse transcriptase promoter mutations, and high uptake of gadoxetic acid in MRI. LC2 is an iCCA-like HCC characterized by expression of the progenitor cell-like trait, tumor protein p53 mutations, and rim arterial-phase hyperenhancement in MRI. LC3 is an HCC-like iCCA, mainly small duct (SD) type, associated with HCC-related etiologic factors. LC4 is further subclassified into LC4-SD and LC4-large duct iCCAs according to the pathological features, which exhibited distinct genetic variations (e.g., *KRAS*,

Abbreviations: *AGR2*, anterior gradient 2; *APOH*, apolipoprotein H; BA, bile acid; BMI, body mass index; CD56, cluster of differentiation 56; cHCC-iCCA, combined HCC and intrahepatic cholangiocarcinoma; CIN, chromosomal instability; CLHCC, cholangiocarcinoma-like HCC; CNA, copy number aberration; *CRP*, C-reactive protein; *CTNNB1*, catenin beta 1; DCNA, DNA copy number–altered gene; DEG, differentially expressed gene; DFS, disease-free survival; ER, endoplasmic reticulum; ESC, embryonic stem cell; EXP, transcriptome expression; FGFR2, FGF receptor 2; *FOXM1*, forkhead box M1; GCA, glycocholate; GDCA, glycodeoxycholate; GRP78, glucose-regulated protein, 78 kDa; HBP, hepatobiliary phase; H&E, hematoxylin-eosin; HNF4A, hepatocyte nuclear factor 4A; *HP*, haptoglobin; iCCA, intrahepatic cholangiocarcinoma; *IDH1/2*, isocitrate dehydrogenase 1/2; LC, liver cancer; LD, large duct; LPC, bipotential liver stem cell; OS, overall survival; rim-APHE, rim arterial-phase hyperenhancement; RNA-seq, RNA sequencing; S100P, S100 calcium binding protein P; SD, small duct; shRNA, short hairpin RNA; SIR-HBP, HBP signal intensity ratio; *SLCO1B1*, solute carrier organic ion transporter family member B1; TCA, taurocholate; TCGA-LIHC, The Cancer Genome Atlas Liver Hepatocellular Carcinoma; TDCA, taurochenodeoxycholate; *TERT*, telomerase reverse transcriptase; *TFF1*, trefoil factor 1; TIGER-LC, Thailand Initiative in Genomics and Expression Research for Liver Cancer; TP53, tumor protein p53; YS, Yonsei University Hospital (cohort).

Youngsic Jeon, So Mee Kwon, and Hyungjin Rhee contributed equally to this work.

© 2022 American Association for the Study of Liver Diseases.

isocitrate dehydrogenase 1/2 mutation, and FGF receptor 2 fusion), stromal type, and prognostic outcomes.

Conclusions: Our integrated view of the molecular spectrum of LCs can identify subtypes associated with transcriptomic, genomic, and radiopathologic features, providing mechanistic insights into heterogeneous LC progression.

INTRODUCTION

HCC and intrahepatic cholangiocarcinoma (iCCA) are the major primary liver cancers (LCs) in adults, derived from hepatocytes and cholangiocytes, respectively. Previously, various molecular subtypes of HCCs and iCCAs have shown distinct clinical and pathological features from different patient cohorts,^[1–3] which could be generally stratified into proliferation and nonproliferation subtypes.^[4] Although HCCs and iCCAs have distinct morphological features and clinical outcomes, there exists an intermediate type of combined HCC and iCCA (cHCC-iCCA).^[5] With this concern, we have demonstrated the presence of overlapping features of HCC and iCCA in cholangiocarcinoma-like HCC (CLHCC),^[6] which it has been suggested is derived from bipotential liver cancer stem cells (LPCs) capable of differentiating into either hepatic or biliary progenitor cells.^[7] In addition, the molecular features of HCCs are related to histopathological features. For example, a scirrhous subtype of HCC is associated with tuberous sclerosis complex subunit 1/2 mutations and stem cell–like expression traits.^[8] Steatohepatic subtype is associated with frequent IL-6/Janus kinase/signal transducer and activator of transcription activation, whereas macrotrabecular-massive subtype is associated with tumor protein p53 (*TP53*) mutation and *FGF19* amplification.^[9] The radiologic features of HCCs are also linked to their molecular characteristics. We have recently reported that the radiological finding of a rim arterial-phase hyperenhancement (rim-APHE) on gadoxetic acid–enhanced MRI is often associated with the stemness gene expression and macrotrabecular pattern in HCC.^[10] iCCAs have also shown enormous molecular heterogeneity, revealing molecular subtypes.^[11,12] Pathologically, iCCAs can be classified into the small duct (SD) and the large duct (LD) types.^[13] SD-iCCA is similar to HCC, showing associations with HCC-related etiologic factors such as viral hepatitis and metabolic syndrome, whereas LD-iCCA is similar to extrahepatic bile duct cancers. Notwithstanding, integrative and systematic analyses of the molecular spectrum with these radiopathologic and clinical features of LCs have not been fully elucidated. In this study, we identified four subtypes representing the molecular spectrum between HCCs and iCCAs, demonstrating their genomic and radiopathologic characteristics. From this, we could recapitulate potential molecular events that drive heterogeneous LC progression.

PATIENTS AND METHODS

Patients and tissue specimens

Patients diagnosed with HCC or iCCA who underwent curative hepatic resection between May 2000 and February 2018 at Severance Hospital, Yonsei University Medical Center, were enrolled. Tissue specimens of liver cancers (YS-LC, $n = 137$), including HCCs ($n = 78$) and iCCAs ($n = 59$), were obtained. The cases of cHCC-iCCA and intermediate carcinoma were excluded according to the criteria of the World Health Organization's *Classification of Digestive System Tumors* (fifth edition).^[14] None of the patients had received preoperative treatment, such as radiation or systemic chemotherapy. The institutional review board of Severance Hospital approved the study (no. 4-2017-0548), and the requirement for informed consent was waived. The overall clinicopathological features of the patients are summarized in Table S1.

Pathology and immunohistochemistry

Whole-section hematoxylin–eosin (H&E)–stained slides of the representative formalin-fixed, paraffin-embedded tissue blocks (>6 tumor cells thick) were analyzed. For iCCAs, pathological types of SD, LD, and undetermined (UD) iCCAs were assessed by immunohistochemical staining of cluster of differentiation 56 (CD56), *N*-cadherin, and S100 calcium binding protein P (S100P).^[15] The fibrous stroma of iCCA was subclassified into mature, intermediate, and immature stroma, as described.^[16] Mature stroma is composed of multilayered mature collagen fibers, whereas immature stroma is composed of myxoid stroma with randomly oriented short keloid-like collagen bundles. Intermediate stroma was defined as keloid-like collagens intermingled with mature stroma.

Radiological analysis

Radiological analysis was performed by an expert abdominal radiologist (H.R.). The uptake of gadoxetic acid in HCC was evaluated in the hepatobiliary phase (HBP) of gadoxetic acid–enhanced MRI. The

high uptake of gadoxetic acid was defined as HCCs showing isointensity or hyperintensity at more than two thirds of the tumor volume. The HBP signal intensity ratio (SIR-HBP) was defined as the ratio of HBP signal intensities in tumors and adjacent non-tumor liver.^[17] The rim-APHE was defined as arterial hyperenhancement most pronounced in the observation periphery (<https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/CT-MRI-LI-RADS-v2018>). The presence of rim-APHE was evaluated in HCC samples ($n = 70$) using dynamic MRI with gadoxetic acid ($n = 40$), dynamic MRI using contrast media other than gadoxetic acid ($n = 13$), or dynamic CT ($n = 17$). The associations of clinicopathological features with the radiologic findings were validated using an independent HCC cohort who underwent hepatic resection at the same institution ($n = 56$; Table S2).

RNA-sequencing data processing and analysis

RNA-sequencing (RNA-seq) profiling was performed using an Illumina HiSeq2500. Chromosomal instability (CIN) of each sample was estimated by calculating the CIN scores using the average expression of the CIN70 signature.^[18] DNA copy number aberrations (CNAs) were estimated by inferring the transcriptome data using inferCNV^[19] with some modifications. The variants identified from RNA-seq data were validated by Sanger sequencing analysis (for details, see Supporting Methods).

Public data analyses

Public data of LC transcriptome were obtained from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) portal sites and Gene Expression Omnibus (GEO) databases. Thailand Initiative in Genomics and Expression Research for Liver Cancer (TIGER-LC) samples with matched DNA copy number profiles and gene expression profiles were used ($n = 152$).^[7] The nontumoral samples and the recurred tumor samples were excluded from the analyses. The LC subtypes were characterized by evaluating the expression of the publicly available molecular classifiers, including the HCC subtype classifiers (i.e., Lee, Woo, TCGA-LIHC, and Hoshida), iCCA subtype classifiers (i.e., Sia, Anderson, and Rhee), and immunotype classifiers (i.e., active immune and exhausted immune). By applying the nearest template prediction algorithm, the subtypes were predicted using the classifiers (false discovery rate [FDR] < 0.05).^[20] The details of the public data sets, gene sets, and subtype classifiers are summarized in Table S3.

Molecular experiments

The details for cell culture, reagents, tissue microarrays, western blotting, chromatin immunoprecipitation (ChIP), cloning, total bile acid assay, site-directed mutagenesis, luciferase transcription factor binding, proliferation, and invasion assays are described in Supporting Methods. The antibodies and the primers used in this study are summarized in Tables S4 and S5, respectively.

Statistical analysis

Statistical analysis was performed using SPSS, version 23.0.1 (SPSS Inc., Chicago, IL) or Rstudio. Clinicopathologic findings between LC subtypes were evaluated using the Student *t* test, the Mann-Whitney U test, the chi-squared test, Fisher's exact test, or the Mantel-Cox log-rank test. Kaplan-Meier plot analyses for overall survival (OS) and disease-free survival (DFS) were performed using a log-rank test.

RESULTS

Molecular spectrum between HCC and iCCA

Transcriptome profiles of LC, including 78 cases of HCC and 59 cases of iCCA, were obtained from the Yonsei University Hospital cohort (YS-LC, $n = 137$). To obtain robust results with an extended sample size, we pooled our data with a previous study, TIGER-LC ($n = 152$), which had the same study design comprising HCCs ($n = 62$) and iCCAs ($n = 90$). By performing a consensus cluster analysis,^[21] we identified four robust subtypes: LC1 ($n = 113$), LC2 ($n = 19$), LC3 ($n = 25$), and LC4 ($n = 132$) (Figure S1; Figure 1A). LC1 and LC2 were mostly comprised of HCCs ($n = 131$, 93.6%), whereas LC3 and LC4 were comprised of iCCAs ($n = 148$, 99.3%). However, when we evaluated the expression levels of the cell marker genes for cholangiocytes and hepatocytes, we observed that LC2 expressed cholangiocyte-like features, whereas LC3 expressed hepatocyte-like features (Figure 1B, top; Figure S2 and Table S3). Principal component analysis also revealed that LC2 and LC3 samples were distributed in the middle between LC1 and LC4, reflecting their intermediate features (Figure 1C). These results suggest that LC2 (HCC, $n = 19$; iCCA, $n = 0$) represents the HCC subtype with iCCA-like feature, whereas LC3 (HCC, $n = 3$; iCCA, $n = 22$) represents the iCCA subtype with HCC-like features. Congruently, expression levels of the CLHCC-like trait and embryonic stem cell (ESC) trait were correlated with the stages of hepatic to cholangiocyte transition from LC1 to LC4 ($p < 0.001$;

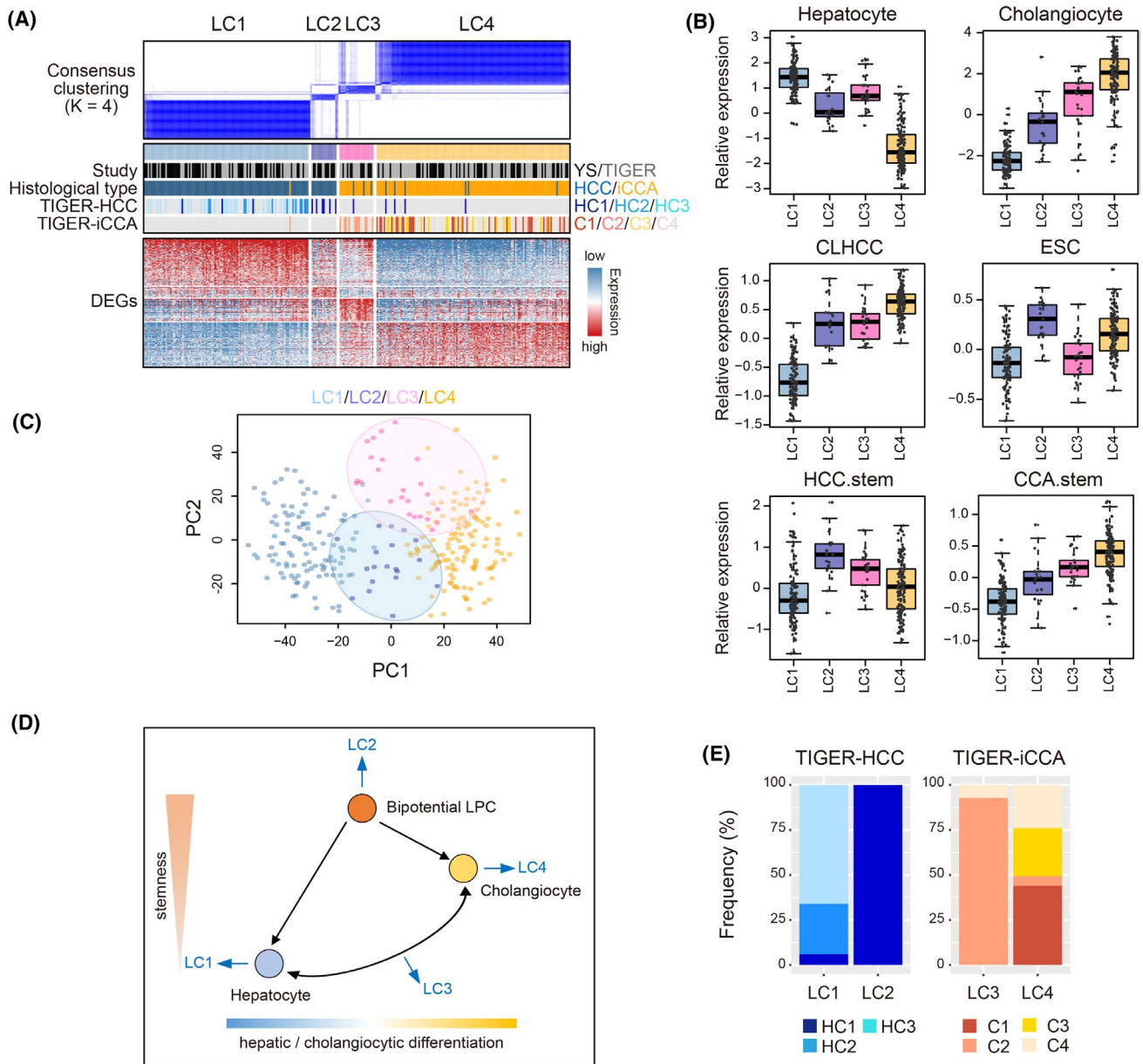


FIGURE 1 The molecular spectrum between HCC and iCCA. (A) Consensus clustering analysis of the integrated data (YS-LC and TIGER-LC) shows the LC subtypes (LC1–LC4, top). Each sample was labeled with the study cohort (YS-LC vs. TIGER-LC), tissue type (HCC vs. iCCA), and the molecular subtypes of TIGER-LC (middle). A heatmap shows the DEGs for each subtype of LC1 ($n = 824$), LC2 ($n = 198$), LC3 ($n = 407$), and LC4 ($n = 785$) (bottom). (B) Boxplots show expression of the cell maker genes across the subtypes, including the gene sets of hepatocyte, cholangiocyte, ESC, and CLHCC and the stem cell signatures of HCC (HCC.stem) and iCCA (CCA.stem). (C) Principal component analysis using the variable genes (median absolute deviation >0.7 , $n = 3,567$) shows that LC2 and LC3 samples are distributed in the middle between LC1 and LC4 samples. (D) A plot shows the putative cell origins of the LC subtypes with expression of the cell origin–related gene signatures. (E) Boxplots show the frequencies of the TIGER-LC subtypes across our LC subtypes for HCC (left) and iCCA (right). PC, principal component

Figure 1B, middle). These findings imply that LC2 and LC3 might be derived from bipotential LPCs. Thus, we evaluated whether LC2 and LC3 express the stemness-related genes, which revealed that LC2 and LC4, but not LC3, harbor LPC-like features. Moreover, we examined the expression of the unipotent progenitor cell traits which were committed to either hepatic lineage cells (i.e., HCC.stem) or biliary lineage cells (i.e., CCA.stem). LC2 exhibited higher expression of HCC.stem

genes than LC1, implying its LPC origin. However, LC3 exhibited lower expression of iCCA.stem genes than LC4 ($p < 0.001$; Figure 1B, bottom), implying its cellular origin from differentiated cells rather than LPCs. LC4 showed higher expression of both cholangiocyte and LPC-like features, which might be due to the LPC-like feature of the cholangiocytes. Hence, we suggest that LC1 and LC4 are derived from mature hepatocytes and cholangiocytes, respectively. LC2 is derived from

LPCs, whereas LC3 is from intermediate and differentiated hepatocytes or cholangiocytes (Figure 1D).

We also demonstrated that several clinicopathological features were significantly associated with the HCC subtypes. LC2-HCCs, compared to LC1-HCCs, showed more aggressive features such as higher alpha-fetoprotein levels (>400 ng/ml, $p = 0.021$), larger tumor size (>5 cm, $p < 0.01$), and frequent microvascular invasion ($p = 0.018$) (Figure S3A). Also, LC4-iCCAs, compared to LC3-iCCAs, demonstrated more aggressive features such as poorer differentiation ($p = 0.047$), large tumor size (>5 cm, $p = 0.021$), and frequent microvascular invasion ($p = 0.016$) (Figure S3B). Correspondingly, Kaplan-Meier survival analysis validated the prognostic predictability of the subtype classifiers in each of four independent HCC data sets (i.e., TCGA-LIHC, GSE4024, GSE113617, and GSE14520) ($p < 0.001$; Figure S4). In addition, comparing our subtypes with the previous subtypes of TIGER-LC, we found that LC2 was similar to the most aggressive TIGER-LC subtype, HC1 (7/7, 100%), whereas LC3 was similar to the less aggressive subtype, C2 (13/14, 92.9%; Figure 1E). We also evaluated other previous molecular classifications associated with HCC aggressiveness (i.e., Lee_A, Woo_C1, TCGA_iCL1, and Hoshida_S1; for details, see Table S3), which consistently support the aggressive phenotype of LC2-HCC compared to LC1-HCC (Figure S5).

Differential DNA copy number alterations and bile acid metabolism between the HCC subtypes

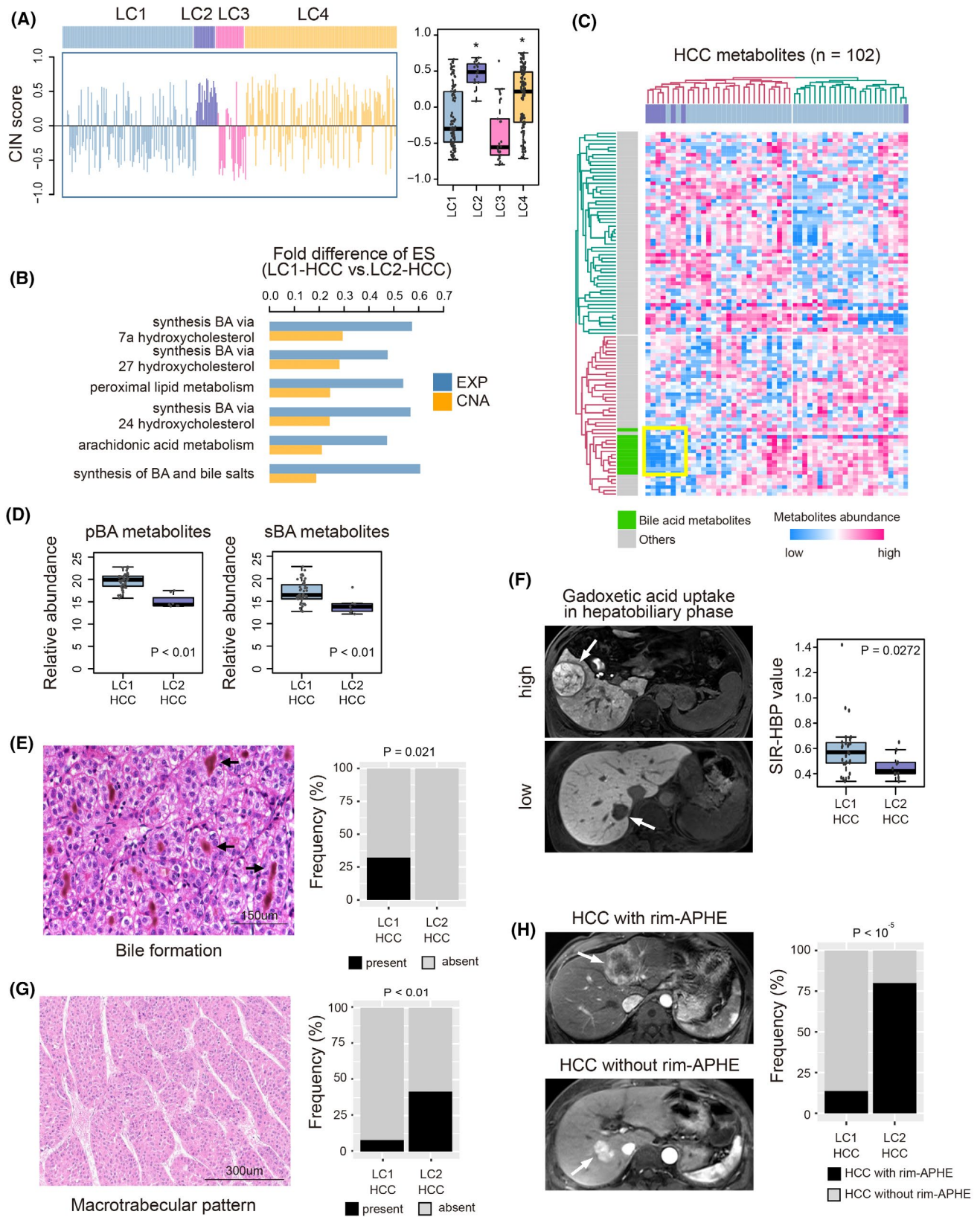
Next, we evaluated CIN across the LC subtypes by calculating CIN scores, as described^[18] (for details, see Supporting Methods). Remarkably, LC subtypes showed distinct CIN scores, indicating that CNAs play crucial roles in developing LC subtypes. LC2 and LC4 exhibited significantly higher CIN scores than LC1 and LC3 (Figure 2A). Thus, we sought to identify key features associated with the distinct CNAs of the HCC subtypes (LC1-HCC vs. LC2-HCC). By performing unsupervised analyses for CNA and transcriptome expression (EXP) profiles, we identified that the bile acid (BA) metabolism-related genes were markedly altered between the HCC subtypes at both the genomic and transcription levels (Figure 2B). To further verify the metabolic alteration in HCC subtypes, we examined the metabolomic profiles of the TIGER-LC data ($n = 102$). Unsupervised hierarchical cluster analysis of the metabolome data revealed that LC1-HCC and LC2-HCC were readily classifiable, indicating distinct metabolic alterations between the HCC subtypes, particularly BA metabolism-related metabolites (Figure 2C). We also found that both primary BA metabolites (cholate, taurocholate [TCA], and glycocholate [GCA]) and secondary BA metabolites (taurochenodeoxycholate [TDCA] and

glycodeoxycholate [GDCA]) were depleted in LC2-HCC rather than LC1-HCC (Figure 2D). This result implies that the HCC subtype differs not only in the synthesis of primary BA but also in the uptake of secondary BA by hepatocytes. The altered BA metabolism was validated by performing microscopic analyses of the YS-HCC specimens ($n = 78$). We found frequent bile formation within the tumor cells of LC1-HCC (20/64) but not in LC2-HCC (0/12, $p = 0.021$; Figure 2E). Thus, we suggest that BA metabolism is a critical feature that distinguishes LC1-HCC and LC2-HCC in all the multi-omic profiles of DNA copy numbers, transcription, and metabolites.

In addition, we examined the radiologic features of HCC subtypes using the HBP uptake of gadoteric acid-enhanced MRI. As gadoteric acid is taken up by the same transporters for BA in HBP images, the HBP signal intensity could be considered a marker for BA uptake.^[22] We could demonstrate that LC1-HCC had significantly higher SIR-HBP than LC2-HCC ($p = 0.027$, Mann-Whitney U test; Figure 2F). On the other hand, LC2-HCC was more frequently associated with macrotrabecular-massive type than LC1-HCC (LC1, 8.06%; LC2, 41.67%; $p < 0.01$, Fisher's exact test; Figure 2G). Previously, the macrotrabecular-massive pattern has been associated with expression of LPC-related genes, rim-APHE, and a worse prognosis.^[10] Consistently, we could demonstrate the higher frequency of rim-APHE in LC2-HCC (83.33%, 10/12) than LC1-HCC (14.29%, 8/56, $p < 10^{-5}$, Fisher's exact test; Figure 2H). Furthermore, we recapitulated the molecular and clinical characteristics of the HCCs with rim-APHE. HCCs with rim-APHE were more prevalent with expression of the LPC-like trait of ESCs ($p < 10^{-4}$) and CLHCC ($p < 10^{-6}$) (Figure 3A) and macrotrabecular-massive type ($p = 0.015$; Figure 3B), showing worse prognostic outcomes (OS, HR, 2.78, $p < 0.001$; DFS, HR, 2.78, $p < 0.001$; Figure 3C). These findings were validated using independent RNA-seq validation HCC data ($n = 56$; see Table S3). HCCs with rim-APHE, compared to HCCs without rim-PAHE, revealed higher expression of ESCs ($p = 0.017$) and CLHCC ($p = 0.013$), frequent macrotrabecular-massive type ($p = 0.036$), and worse prognostic outcomes (extrahepatic metastasis-free survival, HR, 4.55, $p < 0.01$; disease-specific survival, HR, 5.00, $p = 0.011$; OS, HR, 3.13, $p = 0.069$; DFS, HR, 2.08, $p = 0.10$) (Figure 3D–F). Collectively, we suggest that LC1-HCC is a typical HCC characterized by gadoteric acid uptake, whereas LC2-HCC is characterized by LPC-like trait, macrotrabecular-massive type, and rim-APHE.

LC2-HCC expresses anterior gradient 2 by endoplasmic reticulum stress-induced BA depletion

As the DNA copy number-dependent dysregulation of transcription plays potential driver roles in cancer progression, we identified the differentially DNA copy



number-altered genes (DCNAs, $n = 252$, permuted t test $p < 0.001$ and fold difference > 0.1) and the differentially expressed genes (DEGs, $n = 212$, permuted t test $p < 0.001$ and fold difference > 1) between the HCC subtypes. Of these, we identified the CNA-dependent

DEGs ($n = 86$), which had significant correlations between CNAs and their corresponding transcriptional levels (Figure S6). Notably, anterior gradient 2 (*AGR2*) was the most significantly altered gene between the HCC subtypes (Figure 4A), which was confirmed by

FIGURE 2 Altered BA metabolism between HCC subtypes. (A) The CIN scores (left) and LC subtypes (right) are shown. Statistical significance of the subtype difference of LC1 versus LC2 and LC3 versus LC4 are indicated ($*p < 10^{-5}$, Student *t* test). (B) A bar plot shows the differentially enriched gene functions between LC1-HCC and LC2-HCC, which are identified by calculating the enrichment scores of the gene sets (Kyoto Encyclopedia of Genes and Genomes) between LC1-HCC and LC2-HCC from CNA (fold difference >0.1) and EXP (fold difference >0.4) profiles, respectively. (C) Unsupervised clustering analysis of the metabolomic profiles of HCCs ($n = 102$, from TIGER-LC samples) shows the clustered expression of the BA-related metabolites (yellow box). BA metabolites are indicated in the left-side bar (green). (D) Boxplots show the relative abundance of the primary BA and secondary BA metabolites between LC1-HCC and LC2-HCC ($p < 0.01$, Student *t* test). (E) Bile formation in the H&E-stained specimen is indicated (arrows, left), and their frequencies in the HCC subtypes are shown (right). (F) Representative cases showing high or low hepatobiliary uptake in gadoxetic acid-enhanced MRI are shown (left). A boxplot shows the frequency of SIR-HBP (right). (G) Representative H&E-stained images of macrotrabecular-massive type (left) and their frequencies in HCC subtypes (right) are shown. (H) Representative cases of the HCCs with or without rim-APHE in gadoxetic acid-enhanced MRI (left), and the frequencies of the HCCs with rim-APHE are shown in HCC subtypes (right). ES, enrichment score; pBA, primary BA; sBA, secondary BA

immunohistochemical staining, revealing frequent expression of *AGR2* in LC2-HCC (7/12, 58.33 %) but not in LC1-HCC (0/21, $p < 10^{-3}$; Figure 4B).

AGR2 has been reported to promote cancer progression in various cancer types,^[23] which was induced by endoplasmic reticulum (ER) stress.^[24] In support of these findings, we observed that LC2-HCC, compared to LC1-HCC, had higher expression of ER stress-related genes and unfolded protein response-related genes (Figure 4C). We also verified this finding by performing cell experiments. ER stress inducers (i.e., thapsigargin or tunicamycin) could induce *AGR2* expression in various LC cells at the transcription and protein levels (Figure 4D; Figure S7A). Moreover, BA is a chemical chaperone that can reduce intracellular ER stress^[25]; thus, we evaluated whether *AGR2* expression is associated with BA depletion. Treatment with BA metabolites (i.e., GCA, TCA, GDCA, or TDCA) could suppress ER stress-induced *AGR2* expression (Figure S7B). Furthermore, treatment with a heat shock protein 90 inhibitor (PU-H71), which is known to activate ER stress,^[26] could enhance *AGR2* and glucose-regulated protein, 78 kDa (*GRP78*) expression and could be invalidated by cotreatment with BA metabolites such as GCA or TCA (Figure 4E). Cellular ER stress was monitored by examining the phosphorylation levels of PRKR-like ER kinase and eukaryotic translation initiation factor 2a (Figure 4F).

We also demonstrated the clinical significance of *AGR2* expression, showing the shorter OS in patients with HCC and higher *AGR2* expression from independent HCC data sets (TCGA-LIHC and GSE14520; $p < 0.001$; Figure S8A). Also, cell culture experiments successfully demonstrated that *AGR2* enhanced the proliferation and invasion of LC cells (Figure S8B), enhancing expression of the stemness-related genes (e.g., keratin 19 and epithelial cell adhesion molecule; Figure S8C). We also evaluated whether mitogen-activated protein kinase signaling is involved in *AGR2*-induced LC cell progression. We found that overexpression of *AGR2* enhanced the phosphorylation of extracellular signal-regulated kinase (ERK), AKT, and c-Jun N-terminal kinase (JNK) (Figure S8D). Thus, we suggest that *AGR2* expression contributes to the aggressive

phenotype of LC2-HCC potentially through ERK–AKT–JNK activation.

BA metabolism is regulated by hepatocyte nuclear factor 4A-mediated solute carrier organic ion transporter family member B1 expression

Next, we investigated the underlying mechanisms involved in the altered BA metabolism between the HCC subtypes. Indeed, many BA-related genes were CNA-dependent DEGs, such as solute carrier organic ion transporter family member B1 (*SLCO1B1*), BA-CoA:amino acid *N*-acyltransferase (BAAT), solute carrier family 22 member 1 (SLC22A1), solute carrier family 10 member 1 (SLC10A1), and cytochrome P450 family 8 subfamily B member 1 (CYP8B1). Among them, *SLCO1B1* exhibited the highest DNA copy number-dependent correlation of transcription ($r = 0.57$, $p < 10^{-12}$), implying its functional significance (Figure 5A). Indeed, *SLCO1B1* is known as a BA transporter that regulates cellular uptake of BA.^[27] We could demonstrate that short hairpin RNA (shRNA)-mediated knockdown of *SLCO1B1* suppressed BA uptake in liver cancer cells (Figure 5B; Figure S9A).

In addition, we identified 24 putative transcription factors which were predicted to bind the *SLCO1B1* promoter (−2000 to +100 bp from the transcription start site) (Figure 5C). Among them, hepatocyte nuclear factor 4A (*HNF4A*) and forkhead box M1 (*FOXO1*) were the most differentially expressed transcription factors in LC2-HCC and LC1-HCC, respectively. We observed that *HNF4A* target genes such as “*HNF4A* target” ($n = 36$), “SUMI *HNF4A* TARGETS” ($n = 39$), and “*LUCAS_HNF4A* TARGET UP” ($n = 59$) were significantly expressed in LC1-HCC compared to LC2-HCC (Figure S10). Moreover, *HNF4A* expression was closely correlated with the expression levels of *SLCO1B1* and *AGR2* in the six independent HCC data sets, implying that *HNF4A* plays a crucial role in the expression of *SLCO1B1* and *AGR2* (Figure S11). Indeed, *HNF4A* has been shown to regulate BA biosynthesis in the liver.^[28] By performing a luciferase reporter

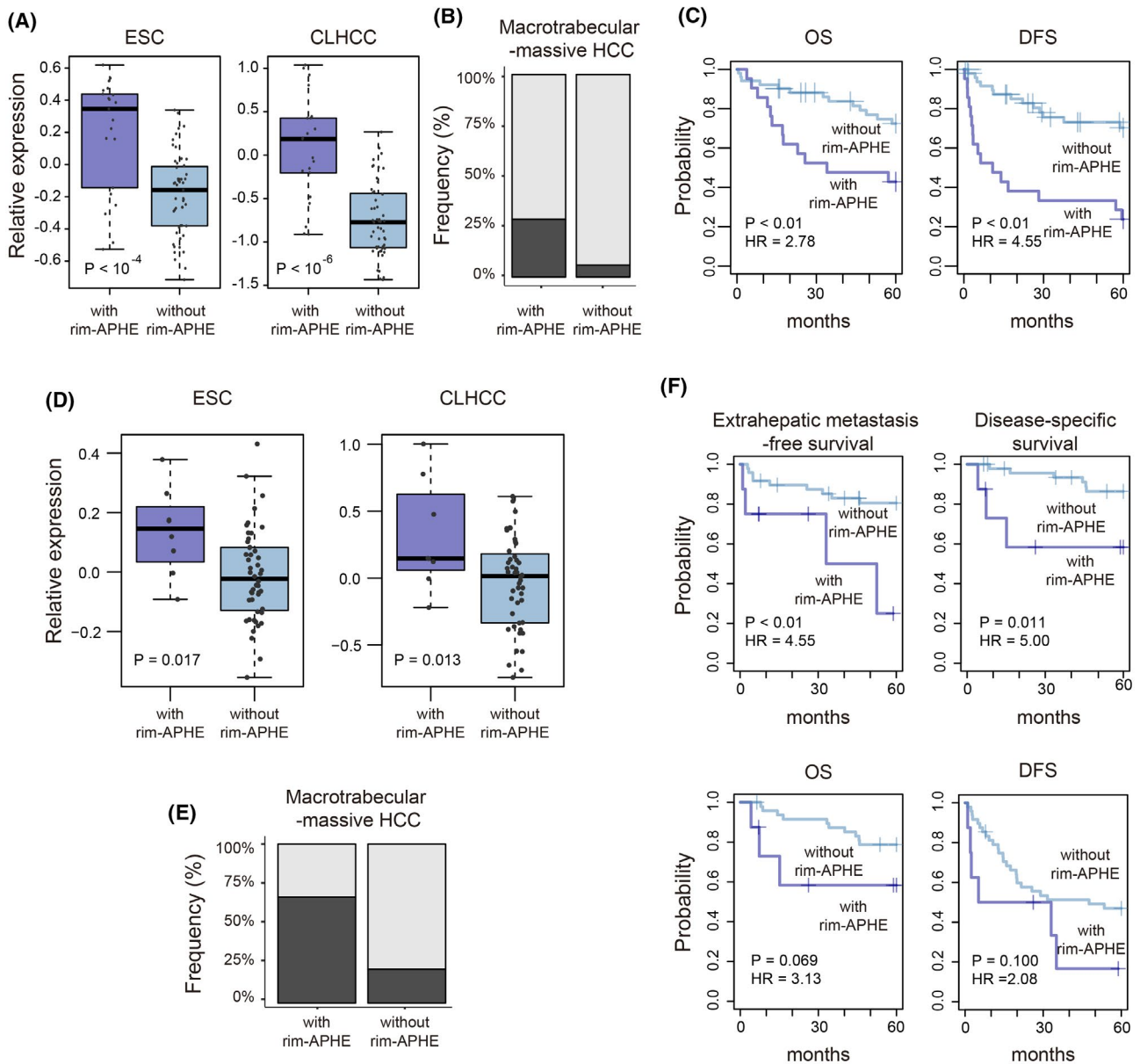


FIGURE 3 HCCs with rim-APHE are associated with LPC-like expression traits, macrotrabecular-massive type, and worse prognosis. (A–C) Gene set enrichment scores of ESC-related and CLHCC-related genes (A), the proportion of the macrotrabecular-massive type (B), and the prognostic outcomes of OS and DFS (C) between patients with HCC with and without rim-APHE (C) are shown in the YS-HCC cohort. (D–F) Gene set enrichment scores of ESC-related and CLHCC-related genes (D), the proportion of the macrotrabecular-massive type (E), and the prognostic outcomes of extrahepatic metastasis-free survival, disease-specific survival, OS, and DFS (F) between patients with HCC with and without rim-APHE are shown in the validation cohort ($n = 56$)

assay using the constructs for the putative *HNF4* binding sites of the *SLCO1B1* promoter, we demonstrated that *HNF4A* binds to the promoter (#2, -39 ~ -30 bp from the transcription start site) and induces *SLCO1B1* transcription (Figure 4D; Figure S12A). ChIP-PCR analysis of the *HNF4* binding sites also confirmed our findings (Figure S12B). The effect of *FOXM1* knock-down on *SLCO1B1* expression was also evaluated, but no significant effect was observed (Figure S9B). Next, we assessed whether *HNF4A*-mediated *SLCO1B1* transcription could affect the expression of *AGR2* and ER stress. Knockdown of *HNF4A* could suppress

SLCO1B1 expression and BA uptake but enhance *AGR2* expression in LC cells (Figure 5E; Figure S13). Furthermore, knockdown of *HNF4* has suppressed gadoxetic acid uptake.^[29] Supporting this, we found that the SIR-HBP scores were significantly correlated with the expression levels of *HNF4* ($r = 0.42$, $p < 10^{-3}$) and *SLCO1B1* ($r = 0.28$, $p < 0.01$; Figure 5F). Taken together, we suggest that the promoter binding of *HNF4A* induces *SLCO1B1* expression, which in turn increases BA uptake and gadoxetic acid uptake, resulting in suppression of ER stress and *AGR2* expression (summarized in Figure 5G).

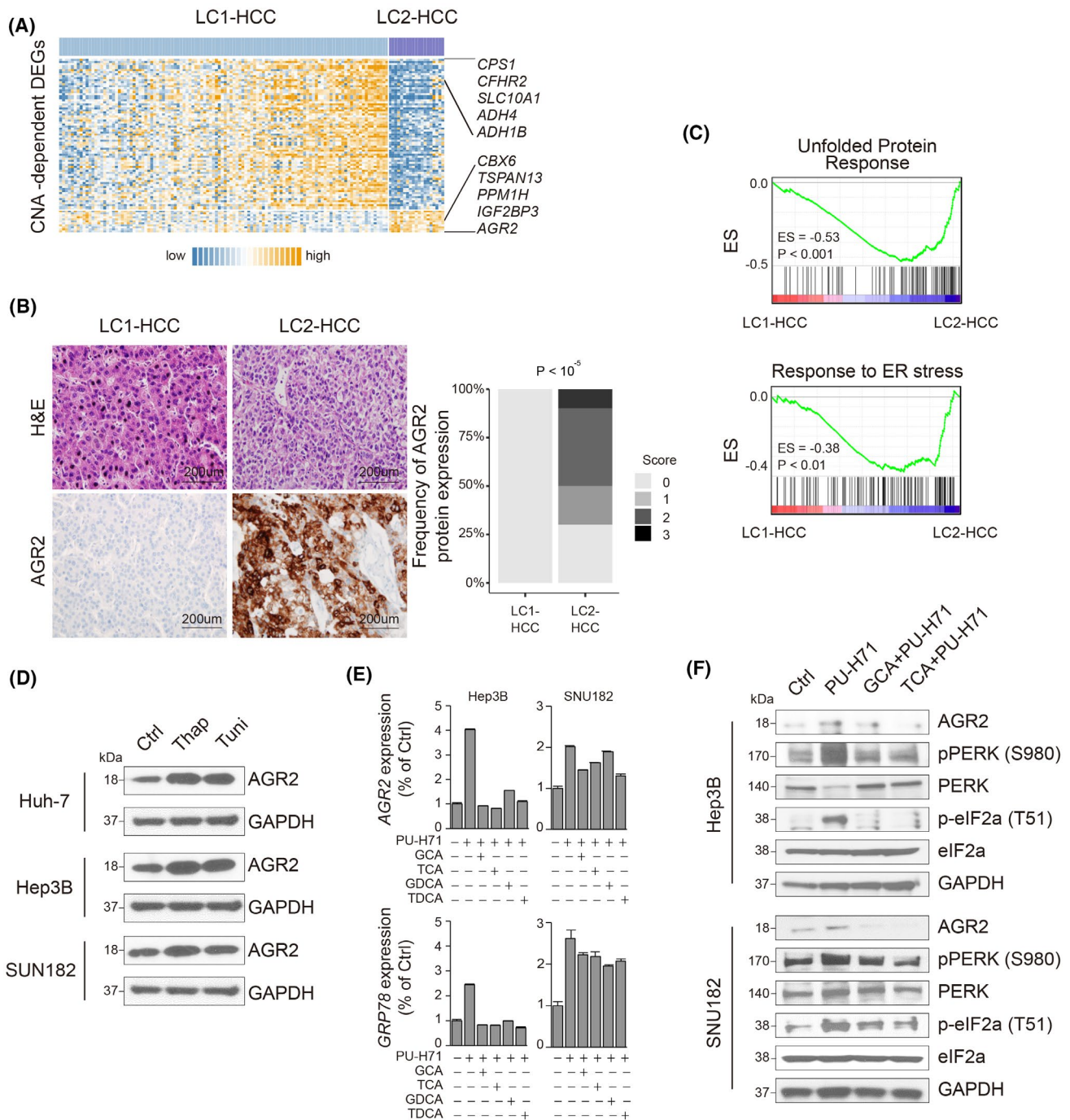
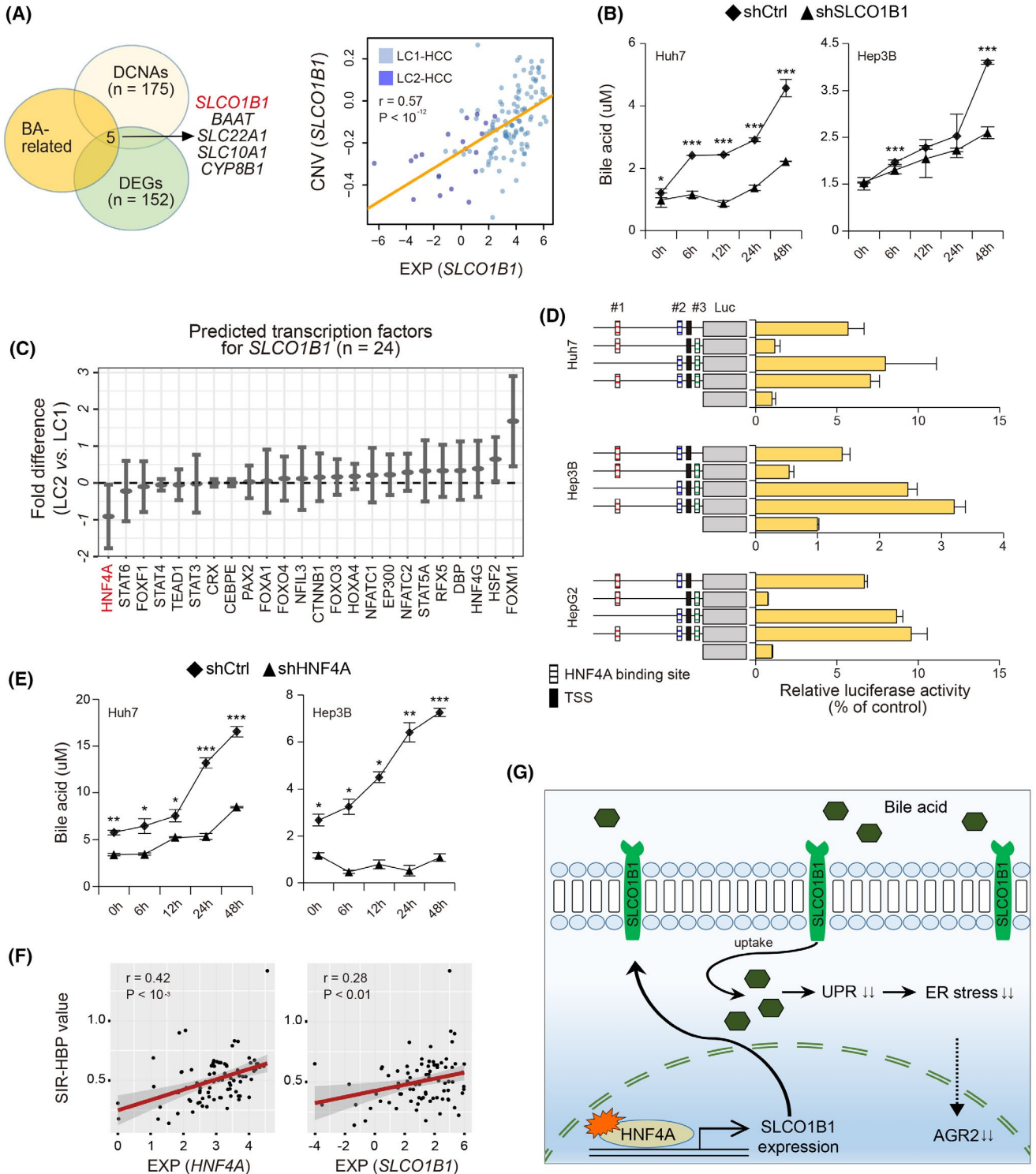


FIGURE 4 BA depletion-induced ER stress facilitates AGR2 expression. (A) A heatmap shows the expression of CNA-dependent genes ($n = 86$, Pearson's correlation coefficient $r > 0.5$, $p < 0.001$) in HCC subtypes. The top-ranked five genes for each subtype are indicated. (B) Histopathological features (top) and immunohistochemical staining for AGR2 (bottom) in HCC subtypes are shown. Frequencies of AGR2 expression in HCC subtypes (right; $p < 10^{-3}$, chi-squared test). (C) Gene set enrichment analyses show enriched expression of the gene sets of "unfolded protein response" and "response to ER stress" in LC2-HCC compared to LC1-HCC. (D) Western blots show expression of AGR2 and glyceraldehyde 3-phosphate dehydrogenase after treatment with thapsigargin (0.5 μM) or tunicamycin (1.0 $\mu\text{g}/\text{ml}$). (E) Bar plots show expression levels of AGR2 (top) and GRP78 (bottom) by PU-H71 treatment with the presence or absence of BAs. (F) Western blots show the expression levels of AGR2, phospho-PERK (S980), PERK, phospho-eIF2 α (T51), and eIF2 α in the Hep3B and SNU182 cell lines treated with PU-H71, GCA, or TCA, respectively. *ADH4/ADH1B*, alcohol dehydrogenase 4/1B; *CBX6*, chromobox 6; *CFHR2*, complement factor H related 2; *CPS1*, carbamoyl phosphate synthase 1; Ctrl, control; ES, enrichment score; *eIF2 α* , eukaryotic translation initiation factor 2 α ; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; *IGF2BP3*, insulin-like growth factor 2 mRNA binding protein 3; p-, phosphorylated; *PERK*, PRKR-like ER kinase; *PPM1H*, protein phosphatase, Mg²⁺/Mn²⁺-dependent 1H; Thap, thapsigargin; *TSPAN13*, tetraspanin 13; Tuni, tunicamycin



Classification of iCCAs with combined transcriptional and pathologic features

Pathologically, iCCAs are classified as SD and LD, distinguished by the differential expression of CD56 and *N*-cadherin in SD-iCCA and musicarmine and S100P in LD-iCCA (Figure 6A). We found that LC3-iCCAs were the SD type except for one case (7/8), whereas LC4-iCCA had both SD type ($n = 25$) and

LD type ($n = 22$, $p < 0.05$; Figure 6B). Considering the histologic and transcriptomic heterogeneity of LC4, we subdivided the LC4-iCCAs into LC4-SD and LC4-LD (Figure 6C). These iCCA subtypes showed substantial alteration of the transcriptome. LC3-SD expressed haptoglobin (*HP*) and apolipoprotein H (*APOH*), LC4-SD expressed C-reactive protein (*CRP*), and LC4-LD expressed trefoil factor 1 (*TFF1*) and *S100P*. Notably, LC3-SD showed more prevalent HBV

FIGURE 5 HNF4A binding to the *SLCO1B1* promoter leads to BA up in HCC. (A) A Venn diagram shows the overlapped genes among the BA-related genes ($n = 88$), DEGs ($n = 212$), and DCNAs ($n = 252$) between LC1-HCCs and LC2-HCCs (left). Correlations between the expression level and the levels of *SLCO1B1* in the HCC subtypes are shown (right). (B) After treatment of the shRNAs for nontarget control shRNA (shCtrl) or *SLCO1B1* (shSLCO1B1) for the indicated times (6, 12, 24, and 48 h), BA levels are measured in the indicated cells ($*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, Student t test). (C) Putative *SLCO1B1* promoter binding transcription factors ($n = 24$) were obtained from GPminer (<http://gpminer.mbc.nctu.edu.tw/>), and expression fold differences between the HCC subtypes are shown. (D) HNF4A binding activities against the putative *SLCO1B1* promoter binding sites (#1, -1275 to ~-1266 bp from the transcription start site; #2, -39 to ~-30 bp from the transcription start site; #3, +6 to ~+15 bp from the transcription start site) are measured by luciferase activity assays in Huh7, Hep3B, and HepG2 cell lines. (E) Huh7 and Hep3B cells treated with control shRNA (shCtrl) or *shHNF4A* for the indicated times (6, 12, 24, and 48 h) and the BA levels are shown ($*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, Student t test). (F) Plots show the correlated expression of *HNF4A* (left) and *SLCO1B1* (right) with SIR-HBP values in patients with HCC ($n = 86$). (G) HNF4A-mediated expression of *SLCO1B1*, resulting in enhanced BA uptake and suppression of ER stress and *AGR2* expression. *CEBPE*, CCAAT enhancer binding protein epsilon; CNV, copy number variation; *CRX*, cone-rod homeobox; *DBP*, D-box binding PAR BZIP transcription factor; *FOXA1/F1/O3/O4*, forkhead box A1/F1/O3/O4; *HOX44*, homeobox A4; *HSF2*, heat shock transcription factor 2; Luc, luciferase; *NFATC1/2*, nuclear factor of activated T cells 1/2; *NFIL3*, nuclear factor IL3 regulated; *PAX2*, paired box 2; *RF5*, regulatory factor X5; *STAT3/4/5A/6*, signal transducer and activator of transcription 3/4/5A/6; *TEAD1*, TEA domain transcription factor 1; TSS, transcription start site; UPR, unfolded protein response

infection (85.7%), higher body mass index (BMI > 25, 42.9%), and diabetes mellitus (57.1%), which might be associated with the HCC-like characteristics of LC3-SD (Figure 6D). In line with this, we observed that LC3-SD significantly expressed metabolism-related, inflammation-related, and immune-related genes, whereas LC4-LD expressed cell proliferation-related and cell cycle-related genes (Figure 6E; $p < 0.05$). LC4-SD showed intermediate phenotypes between LC3-SD and LC4-LD.

In addition, we evaluated carefully whether LC3-SD and LC4-SD are different histopathologically. Previously, fibrous tumor stroma were classified into mature, intermediate, and immature types according to their stromal features (see Patients and Methods). Immature fibrous stroma has been suggested to facilitate tumor growth and progression as well as immunosuppression, showing worse prognostic outcomes of iCCAs.^[30] In support of this, we observed a significant difference between the fibrous stroma types of LC3-SD and LC4-SD (Figure 6F). LC3-SD was comprised of mature stroma (3/7) and intermediate stroma (4/7), while LC4-SD had frequent immature stroma (20/26) but no mature stroma ($p < 10^{-4}$). LC4-LD showed more prevalent immature stroma (20/21). Furthermore, we observed that LC4-SD had higher CIN scores ($p < 0.01$; Figure S14A) and more frequent exhausted immunotype than LC3-SD ($p = 0.042$, chi-squared test; Figure S14B). These findings consistently indicate that LC4-SD, compared to LC3-SD, had more aggressive tumor behaviors, exhibiting prevalent immature stroma and exhausted immunotype features. Indeed, we could demonstrate the distinct prognostic outcomes among the iCCA subtypes, revealing the worst prognosis of LC4-LD, intermediate prognosis of LC4-SD, and the most favorable prognosis of LC3-SD iCCAs (OS, HR, 24.3, $p < 10^{-5}$; DFS, HR, 19.0, $p < 10^{-4}$; Figure 6G). We suggest that the iCCA subtypes based on the combined transcriptomic and pathological features can represent distinct biological and clinical features of iCCAs.

LC subtypes show distinct mutation profiles

Next, we compared the mutation profiles of the LC subtypes. Overall, tumor mutation burdens differed across the subtypes, showing more frequent mutations in LC2 (average frequencies for nonsynonymous and synonymous variants, 71.1 and 50.3) and LC4 (64.7 and 48.8) compared to those of LC1 (57.2 and 44.1) and LC3 (59.1 and 42.6) (Figure 7A), which might be related to their different chromosomal stability across the subtypes. Interestingly, we found that telomerase reverse transcriptase (*TERT*) promoter mutations were frequent in LC1-HCCs (19/62, 30.6%) but not in LC2-HCCs (0/12, 0%), whereas *TP53* mutations were more frequent in LC2-HCCs (6/12, 50%) than LC1-HCCs (8/62, 12.9%) (Figure 7B,C). On the other hand, the frequencies of catenin beta 1 (*CTNNB1*) mutations were not significantly different between LC1-HCCs (10/62, 16.12%) and LC2-HCCs (2/12, 16.7%). The mutations of *TP53* (12/14) and *CTNNB1* (9/12) were validated by Sanger sequencing (Figure S15A). In particular, we noticed that the higher *TP53* mutation frequency in LC2 might be related to the radiologic rim-APHE feature. Therefore, we further analyzed the *TP53* mutations with an extended data set ($n = 137$) by adding the 56 validation data and the nine more HCC cases with rim-APHE, which confirmed the more prevalent *TP53* mutations in the HCCs with rim-APHE (14/38, 36.8%) than the HCCs without rim-APHE (13/99, 13.1%, $p < 0.01$; Figure 7D). Thus, we suggest that rim-APHE is associated with *TP53* mutations, although further extended studies are required.

In mutation analysis, we observed that isocitrate dehydrogenase 1/2 (*IDH1/2*) mutations were more prevalent in SD-iCCAs (LC3-SD, 2/7, 28.6%; LC4-SD, 8/25, 32.0%), whereas *KRAS* mutations were more prevalent in LD-iCCAs (3/22, 13.6%), as reported^[13] (see Figure 7C). Interestingly, we found four cases of FGF receptor 2 (*FGFR2*) fusion transcripts exclusively in LC4-SD, including the previously known fusions of *FGFR2*-adenosylhomocysteinase like 1 and

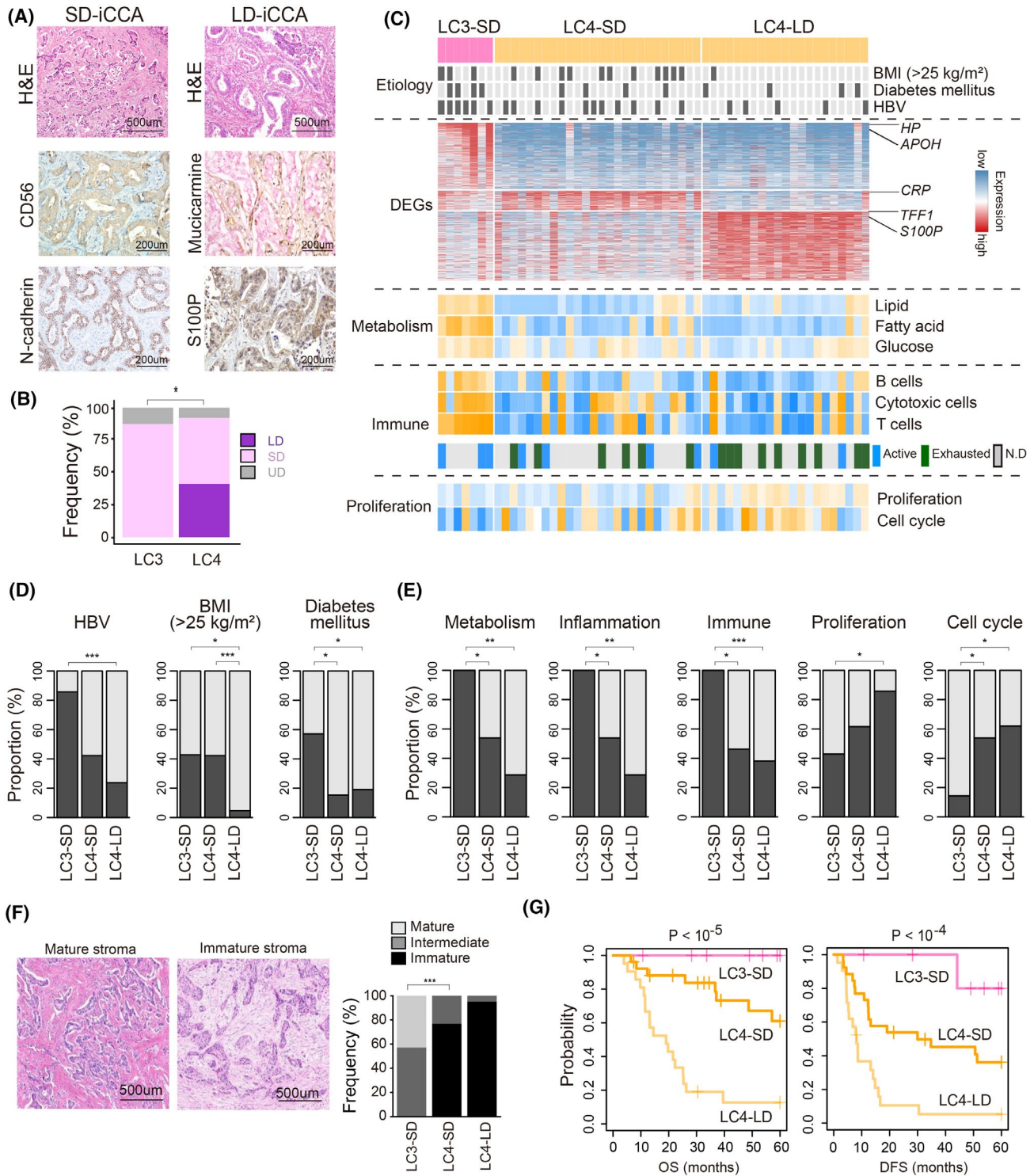


FIGURE 6 Subclassification of iCCAs with combined pathological and transcriptomic features. (A) Histological images of SD-iCCA and LD-iCCA (H&E, top). Mucin and S100P expression in LD-iCCA and CD56 and N-cadherin in SD-iCCA are shown (bottom). (B) Frequencies of LD-iCCA, SD-iCCA, and UD-iCCA in iCCA subtypes are shown (chi-squared test). (C) A heatmap and bars show the clinical etiological factors (first panel, $p < 0.05$), DEGs (second panel, $n = 988$), and the enriched expression scores for the gene sets for the functions (third panel, i.e., metabolism-related, immune-related, and proliferation-related genes), immune cell types (i.e., B cells, cytotoxic cells, and T cells), and immunotypes (i.e., active, exhausted, and not determined, fourth panel), and the proliferation-related genes (fifth panel). $FDR \geq 0.05$. (D,E) Bar plots show the frequencies of the etiological factors of HBV, BMI (>25 kg/m²), and diabetes mellitus (D) and the enriched expression scores for cell metabolism-related, inflammation-related, immune-related, cell proliferation-related, and cell cycle-related genes across the iCCA subtypes (E) ($*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, Fisher's exact test). (F) Pathologic images of mature and immature stroma in iCCA are shown (left). Frequencies of the stroma types are shown in iCCA subtypes (right, chi-squared test). (G) Kaplan-Meier plot analyses for OS and DFS are shown in iCCA subtypes. N.D., not determined

FGFR2–WW domain containing adaptor with coiled-coil^[31] and the new fusions of *FGFR2*–catalase and *FGFR2*–RNA binding Fox 1 homolog 2 (Figure 7E), which we validated by Sanger sequencing (Figure S15B). In addition, when we evaluated the pathway mutations, LC1 and LC2 showed frequent mutations of the WNT pathway (adenomatous polyposis coli, *CTNNB1*, E1A binding protein P300 [EP300], Axin-1, and transcription factor 7 like 2), and LC2-HCC had frequent mutations of DNA repair-related pathway (*TP53*, ataxia telangiectasia and Rad3-related, and partner and localizer of breast cancer type 2) (Figure 7F). SD-iCCAs were frequently mutated with the epigenetic regulators (>20%; cAMP responsive element binding protein binding protein, *IDH1*, and *IDH2*), whereas the LD-iCCAs (LC4-LD) were frequently mutated with the *KRAS* pathway (36.4%; *KRAS*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha [*PI3KCA*], B-raf proto-oncogene, serine/threonine kinase 11 [*STK11*], *NRAS*, and *MET*), implying that the distinct mutation profiles may contribute to the development and progression of LC subtypes.

DISCUSSION

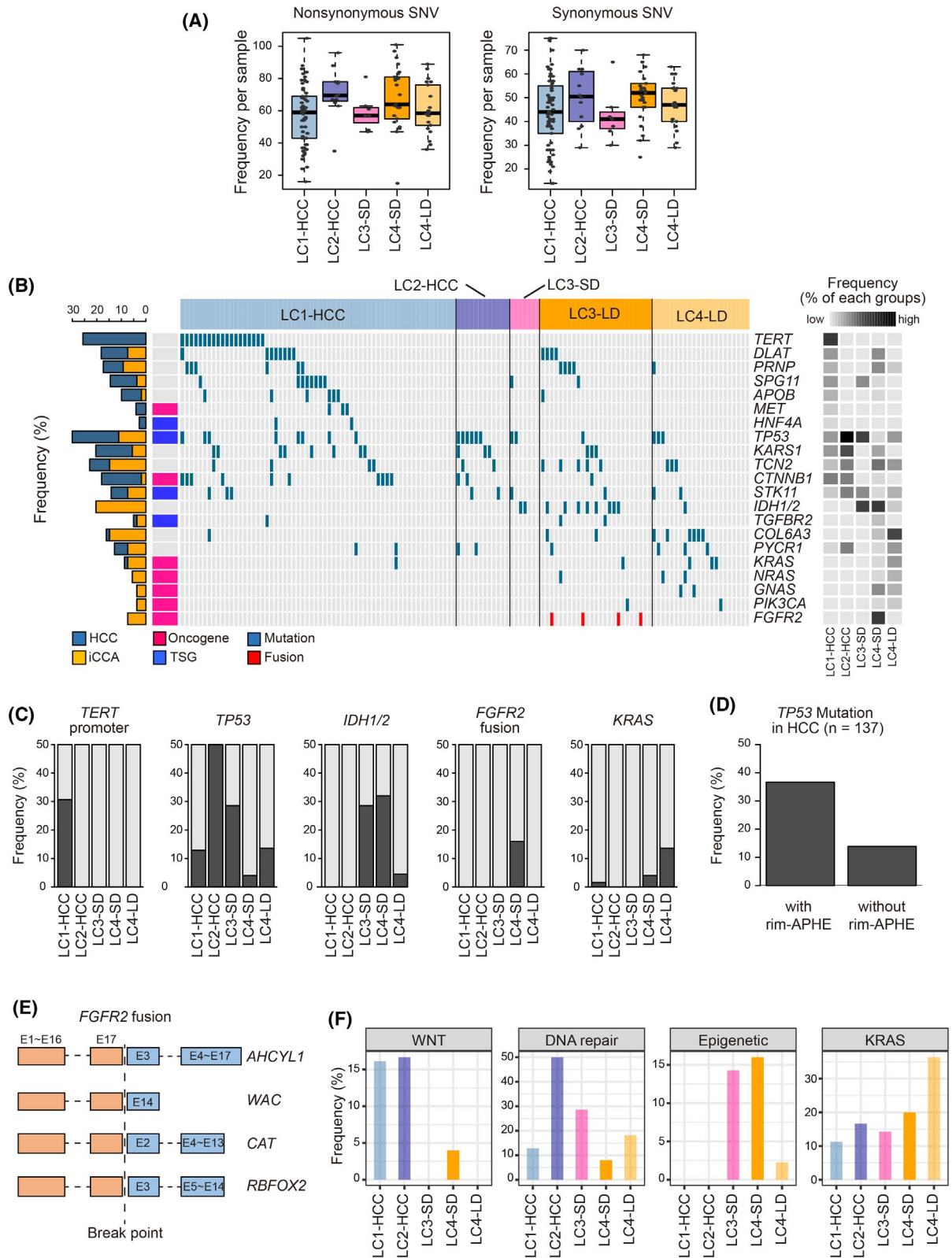
HCCs and iCCAs are derived from diverse differentiation stages of LPCs, hepatocytes, and cholangiocytes, conferring heterogeneous molecular features and clinical outcomes of the tumors. Thus, the concept of the continuous molecular spectrum between HCCs and iCCAs might be beneficial for understanding the molecular heterogeneity and underlying pathobiological mechanisms. In this study, we performed integrative analyses of the LC transcriptome with the perspective of radiopathologic features, revealing the intermediate molecular spectrum of LC subtypes (summarized in Figure 8). LC1 is a typical HCC, whereas LC2 is an atypical HCC with LPC-like trait and aggressive behavior, implying a bipotential LPC origin. Moreover, LC2 showed frequent *TP53* mutation, suggesting that LC2 originates from dedifferentiated hepatocytes because the loss of *TP53* can facilitate dedifferentiation of mature hepatocytes into nestin-positive progenitor-like cells.^[32] Similarly, cHCC-iCCA has been shown to express stem cell-like traits and nestin with a higher frequency of *TP53* mutations.^[5] We could demonstrate the higher expression of nestin in LC2 than in LC1, implying that LC2 originated from dedifferentiated LPC cells ($p < 0.001$; Figure S16).

In-depth analysis of the LC subtypes revealed that LC1-HCC was characterized by active BA metabolism and higher uptake of gadoxetic acid, whereas LC2-HCC was characterized by rim-APHE in MRI and BA depletion. It has been shown that uptake of gadoxetic acid through the BA transporter *SLCO1B3* leads to gadoxetic acid enhancement on MRI.^[33] LC1-HCC,

compared to LC2-HCC, had higher DNA copy number gains and concomitant transcription of *SLCO1B1* ($p < 0.001$) than *SLCO1B3* ($p < 0.05$; Figure S17). Moreover, we demonstrated that HNF4A binding to the promoter of *SLCO1B1* plays a critical role in the altered BA metabolism of HCCs (see Figure 4). Consequently, the depleted BA metabolism could increase ER stress and enhance *AGR2* expression, promoting an aggressive HCC phenotype. However, we also demonstrated that the CNA gains of *AGR2* may contribute to *AGR2* expression in the aggressive LC2-HCC.

Notably, the LC1-HCCs were associated with Hoshida's S3 subclass (46.4%), which was reported to have frequent *CTNNB1* mutations.^[3] *CTNNB1*-mutated HCCs were also reported to show frequent gadoxetic acid uptake.^[34] Concerning this, we could demonstrate that the *CTNNB1*-mutated tumors frequently had higher SIR-HBP scores (Figure S18); however, LC1-HCCs did not show frequent *CTNNB1* mutations, although they had high SIR-HBP scores. This discrepancy may be due to ethnic or etiological differences (HBV-related) of the HCC cohorts among the studies. Indeed, our Korean patients with HCC had a lower frequency of *CTNNB1* mutations than the previous studies (25%–40%).^(35,36)

LC3 is an HCC-like iCCA that expressed both hepatocyte and cholangiocyte expression traits but lower expression of progenitor-like traits; hence, the intermediate feature of LC3 might be derived from differentiated hepatocytes or cholangiocytes. Moreover, previous experimental studies have demonstrated that iCCAs might originate from mature hepatocytes, indicating cholangiocyte-to-hepatocyte transdifferentiation.^[37] This may imply that LC3 is derived from transdifferentiated cells. In addition, we further classified iCCAs into three subtypes (i.e., LC3-SD, LC4-SD, and LC4-LD) based on the combined transcriptomic and pathologic features, revealing phenotypic transition across the subtypes. LC3-SD showed HCC-like clinical features, showing prevalent HCC-etiological factors such as HBV infection and metabolic syndrome (see Figure 6). Notably, both LC3-SD and LC4-SD are pathologically SD type, but they showed differences in the expression of the immune metabolism-related genes and chromosomal stability. In addition, LC3-SD was frequent with the active immunotype and the mature stroma, whereas LC4-LD was frequent with the exhausted immunotype and the immature stroma. This finding implies that the immunotypes were associated with their stroma types, although further elucidation might be required. In mutation analysis, we found *FGFR2* fusions exclusively in LC4-SD, whereas the *KRAS* pathway mutations were more frequent in LC4-LD. *FGFR2* fusion has recently received great attention because its clinical utility as a therapeutic target has been suggested.^[38] As the *FGFR2* fusions and *KRAS* mutations are mutually exclusive,^[39] our results suggest that the exclusive *FGFR2* fusions and *KRAS* mutations may be attributable to the pathologic subtypes.



We classified the LCs using the integrated data set, which could be affected by biased batch effects because of the different cohorts and data platforms. Thus, we carefully integrated the two data by removing the batch effects, taking advantage of the integrative

analysis to identify the robust results regardless of cohort composition or data platforms. We also reevaluated our classification using the YS-LC data alone to rule out possible batch effects or sample biases by the data integration. We could reproduce the four subtypes,

FIGURE 7 LC subtypes have distinct mutation profiles. (A) Boxplots show the frequencies of nonsynonymous (left) and synonymous (right) mutations across LC subtypes (B) A heatmap shows the differentially mutated genes across the LC subtypes (middle). Mutation frequencies in HCCs and iCCAs are indicated (left). Mutation rates in each subtype are shown in a scale-colored heatmap (right). (C) Bar plots show the different mutation frequencies of the *TERT* promoter, *TP53*, *IDH1/2*, *FGFR2* fusion, and *KRAS* across the LC subtypes. (D) The *TP53* mutation frequency in HCCs with rim-APHE is shown. (E) *FGF2R* fusions with four different genes are schematically illustrated. (F) Frequencies of the pathway mutations for WNT, epigenetic regulators, DNA repair, and *KRAS* are shown in LC subtypes. *AHCYL1*, adenosylhomocysteinase like 1; *APOB*, apolipoprotein B; *CAT*, catalase; *COL6A3*, collagen type VI alpha 3 chain; *DLAT*, dihydrolipoamide S-acetyltransferase; *GNAS*, G protein subunit alpha S; *KARS1*, lysyl-transfer RNA synthetase 1; *PRNP*, prion protein; *PYCR1*, pyrroline-5-carboxylate reductase 1; *RBFOX2*, RNA binding Fox 1 homolog 2; SNV, single nucleotide variant; *SPG11*, spastic paraplegia 11; *TCN2*, transcobalamin 2; *TSG*, tumor suppressor gene; *TGFB2*, TGF-beta receptor 2; *WAC*, WW domain containing adaptor with coiled-coil

Subtype	LC1	LC2	LC3-SD	LC4-SD	LC4-LD
Tissue characteristics	HCC	iCCA-like HCC	HCC-like iCCA	iCCA	iCCA
Cells of origin	Hepatocyte	LPC	Hepatocytes or cholangiocytes	Cholangiocyte	Cholangiocyte
Radiology	Gadoxetic acid -uptake	rim-APHE			
Histo-pathology		Macrotrabecular-massive	SD mature stroma	SD immature stroma	LD immature stroma
Metabolism	Bile acid		Glucose/lipid		
Immune			Activated		Exhausted
ER stress					
CIN					
Gene expression	<i>HNF4A</i> <i>SLCO1B1</i>	<i>AGR2</i>	<i>HP</i> <i>APOH</i>	<i>CRP</i>	<i>TFF1</i> <i>S100P</i>
Mutation	<i>TERT</i> promoter	<i>TP53</i>	<i>IDH1/2</i>	<i>IDH1/2</i> <i>FGFR2</i> fusion	<i>KRAS</i>
Prognosis		Unfavorable			Unfavorable

FIGURE 8 Summary of genomic and radiopathologic features of the LC subtypes

revealing the distinct subtype-dependent phenotypes (Figure S19). Therefore, we suggest that our integrative analyses are not likely to be false or to have biased observations.

In conclusion, we suggest that the integrated view of the molecular spectrum between HCCs and iCCAs can identify molecular subtypes representing distinct genomic and radiopathologic features, providing pathobiological insights into the heterogeneous progression of LC.

CONFLICT OF INTERESTS

Nothing to disclose.

AUTHOR CONTRIBUTIONS

Youngsic Jeon performed data analyses and molecular experiments. So Mee Kwon performed data analyses. Hyungjin Rhee performed radiological data analyses. Jeong Eun Yoo and Taek Chung analyzed immunohistochemical images and clinicopathologic data. Hyun Goo Woo and Young Nyun Park developed the study design, performed data analyses, and supervised and funded the study. All authors contributed to writing the manuscript, reviewed the results, and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

The transcriptomic profile data are available in the GEO database (<http://www.ncbi.nlm.nih.gov/projects/geo>) under accession number GSE179443.

ORCID

Youngsic Jeon  <https://orcid.org/0000-0002-5673-5036>
 So Mee Kwon  <https://orcid.org/0000-0001-8017-9650>
 Hyungjin Rhee  <https://orcid.org/0000-0001-7759-4458>
 Jeong Eun Yoo  <https://orcid.org/0000-0001-9990-279X>
 Taek Chung  <https://orcid.org/0000-0001-7567-0680>
 Hyun Goo Woo  <https://orcid.org/0000-0002-0916-893X>
 Young Nyun Park  <https://orcid.org/0000-0003-0357-7967>

REFERENCES

- Chaisaingmongkol J, Budhu A, Dang H, Rabibhadana S, Pupacdi B, Kwon SM, et al. Common molecular subtypes among Asian hepatocellular carcinoma and cholangiocarcinoma. *Cancer Cell*. 2017;32:57–70.e3.
- Hoshida Y, Nijman SMB, Kobayashi M, Chan JA, Brunet J-P, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res*. 2009;69:7385–92.
- Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell*. 2017;169:1327–41.e23.
- Calderaro J, Ziol M, Paradis V, Zucman-Rossi J. Molecular and histological correlations in liver cancer. *J Hepatol*. 2019;71:616–30.
- Xue R, Chen LU, Zhang C, Fujita M, Li R, Yan S-M, et al. Genomic and transcriptomic profiling of combined hepatocellular and intrahepatic cholangiocarcinoma reveals distinct molecular subtypes. *Cancer Cell*. 2019;35:932–47.e8.
- Woo HG, Lee J-H, Yoon J-H, Kim CY, Lee H-S, Jang JJ, et al. Identification of a cholangiocarcinoma-like gene expression trait in hepatocellular carcinoma. *Cancer Res*. 2010;70:3034–41.
- Mishra L, Banker T, Murray J, Byers S, Thenappan A, He AR, et al. Liver stem cells and hepatocellular carcinoma. *Hepatology*. 2009;49:318–29.
- Seok JY, Na DC, Woo HG, Roncalli M, Kwon SM, Yoo JE, et al. A fibrous stromal component in hepatocellular carcinoma reveals a cholangiocarcinoma-like gene expression trait and epithelial–mesenchymal transition. *Hepatology*. 2012;55:1776–86.
- Calderaro J, Couchy G, Imbeaud S, Amaddeo G, Letouze E, Blanc J-F, et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. *J Hepatol*. 2017;67:727–38.
- Rhee H, An C, Kim HY, Yoo JE, Park YN, Kim MJ. Hepatocellular carcinoma with irregular rim-like arterial phase hyperenhancement: more aggressive pathologic features. *Liver Cancer*. 2019;8:24–40.
- Sia D, Hoshida Y, Villanueva A, Roayaie S, Ferrer J, Tabak B, et al. Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. *Gastroenterology*. 2013;144:829–40.
- Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov*. 2017;7:1116–35.
- Banales JM, Marin JGG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nat Rev Gastroenterol Hepatol*. 2020;17:557–88.
- Torbenson MSNI, Park YN, Roncalli M, Sakamoto M. Hepatocellular carcinoma. In: WHO classification of tumours of the digestive system. 5th ed. Lyon: International Agency for Research on Cancer; 2019. pp. 229–39.
- Chung T, Rhee H, Nahm JH, Jeon Y, Yoo JE, Kim Y-J, et al. Clinicopathological characteristics of intrahepatic cholangiocarcinoma according to gross morphologic type: cholangiolocellular differentiation traits and inflammation- and proliferation-phenotypes. *HPB (Oxford)*. 2020;22(6):864–73.
- Zhang XF, Dong M, Pan YH, Chen JN, Huang XQ, Jin Y, et al. Expression pattern of cancer-associated fibroblast and its clinical relevance in intrahepatic cholangiocarcinoma. *Hum Pathol*. 2017;65:92–100.
- Kim JY, Kim MJ, Kim KA, Jeong HT, Park YN. Hyperintense HCC on hepatobiliary phase images of gadoteric acid-enhanced MRI: correlation with clinical and pathological features. *Eur J Radiol*. 2012;81:3877–82.
- 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;38:1043–8.
- Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014;344:1396–401.
- Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. *PLoS One*. 2010;5:e15543.
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26:1572–3.
- Halilbasic E, Claudel T, Trauner M. Bile acid transporters and regulatory nuclear receptors in the liver and beyond. *J Hepatol*. 2013;58:155–68.
- Tiemann K, Garri C, Lee SB, Malhi PD, Park M, Alvarez RM, et al. Loss of ER retention motif of AGR2 can impact mTORC signaling and promote cancer metastasis. *Oncogene*. 2019;38:3003–18.
- Dumartin L, Alrawashdeh W, Trabulo SM, Radon TP, Steiger K, Feakins RM, et al. ER stress protein AGR2 precedes and is involved in the regulation of pancreatic cancer initiation. *Oncogene*. 2017;36:3094–103.
- de Almeida SF, Picarote G, Fleming JV, Carmo-Fonseca M, Azevedo JE, de Sousa M. Chemical chaperones reduce endoplasmic reticulum stress and prevent mutant HFE aggregate formation. *J Biol Chem*. 2007;282:27905–12.
- Davenport EL, Moore HE, Dunlop AS, Sharp SY, Workman P, Morgan GJ, et al. Heat shock protein inhibition is associated with activation of the unfolded protein response pathway in myeloma plasma cells. *Blood*. 2007;110:2641–9.
- Ieiri I, Suwannakul S, Maeda K, Uchimarum H, Hashimoto K, Kimura M, et al. SLCO1B1 (OATP1B1, an uptake transporter) and ABCG2 (BCRP, an efflux transporter) variant alleles and pharmacokinetics of pitavastatin in healthy volunteers. *Clin Pharmacol Ther*. 2007;82:541–7.
- Lau HH, Ng NHJ, Loo LSW, Jasmen JB, Teo AKK. The molecular functions of hepatocyte nuclear factors—in and beyond the liver. *J Hepatol*. 2018;68:1033–48.
- Yamashita T, Kitao A, Matsui O, Hayashi T, Nio K, Kondo M, et al. Gd-E0B-DTPA-enhanced magnetic resonance imaging and alpha-fetoprotein predict prognosis of early-stage hepatocellular carcinoma. *Hepatology*. 2014;60:1674–85.
- De Jaeghere EA, Denys HG, De Wever O. Fibroblasts fuel immune escape in the tumor microenvironment. *Trends Cancer*. 2019;5:704–23.
- Silverman IM, Hollebecque A, Friboulet L, Owens S, Newton RC, Zhen H, et al. Clinicogenomic analysis of FGFR2-rearranged cholangiocarcinoma identifies correlates of

- response and mechanisms of resistance to pemigatinib. *Cancer Discov.* 2021;11:326–39. <https://doi.org/10.1158/2159-8290.cd-20-0766>
32. Tschaharganeh D, Xue W, Calvisi D, Evert M, Michurina T, Dow L, et al. p53-dependent nestin regulation links tumor suppression to cellular plasticity in liver cancer. *Cell.* 2014;158:579–92.
 33. Ueno A, Masugi Y, Yamazaki K, Komuta M, Effendi K, Tanami Y, et al. OATP1B3 expression is strongly associated with Wnt/ β -catenin signalling and represents the transporter of gadoxetic acid in hepatocellular carcinoma. *J Hepatol.* 2014;61:1080–7.
 34. Kitao A, Matsui O, Yoneda N, Kozaka K, Kobayashi S, Sanada J, et al. Hepatocellular carcinoma with β -catenin mutation: imaging and pathologic characteristics. *Radiology.* 2015;275:708–17.
 35. Zucman-Rossi J, Villanueva A, Nault J-C, Llovet JM. Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology.* 2015;149:1226–39.e4.
 36. Lombardo D, Saitta C, Giosa D, Di Tocco FC, Musolino C, Caminiti G, et al. Frequency of somatic mutations in *TERT* promoter, *TP53* and *CTNNB1* genes in patients with hepatocellular carcinoma from southern Italy. *Oncol Lett.* 2020;19:2368–74.
 37. Fan B, Malato Y, Calvisi DF, Naqvi S, Razumilava N, Ribback S, et al. Cholangiocarcinomas can originate from hepatocytes in mice. *J Clin Invest.* 2012;122:2911–5.
 38. Cristinziano G, Porru M, Lamberti D, Buglioni S, Rollo F, Amoreo CA, et al. FGFR2 fusion proteins drive oncogenic transformation of mouse liver organoids towards cholangiocarcinoma. *J Hepatol.* 2021;75:351–62.
 39. Arai Y, Totoki Y, Hosoda F, Shirota T, Hama N, Nakamura H, et al. Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology.* 2014;59:1427–34.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Jeon Y, Kwon SM, Rhee H, Yoo JE, Chung T, Woo HG, et al. Molecular and radiopathologic spectrum between HCC and intrahepatic cholangiocarcinoma. *Hepatology.* 2022;00:1–17. <https://doi.org/10.1002/hep.32397>