Carcinogenesis, 2021, Vol. 42, No. 5, 672–684

doi:10.1093/carcin/bgab014
Advance Access publication 22 February 2021
Original Article

ORIGINAL ARTICLE

Oncogenic transcriptomic profile is sustained in the liver after the eradication of the hepatitis C virus

Haruhiko Takeda, Atsushi Takai*, Eriko Iguchi, Masako Mishima, Soichi Arasawa, Ken Kumagai, Yuji Eso, Takahiro Shimizu, Ken Takahashi, Yoshihide Ueda[1], Kojiro Taura[2], Etsuro Hatano[2,3], Hiroko Iijima[4], Haruyo Aoyagi[5], Hideki Aizaki[5], Hiroyuki Marusawa[6], Takaji Wakita[5] and Hiroshi Seno

Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [1]Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kobe University, Kobe, Japan; [2]Division of Hepato-Biliary-Pancreatic and Transplant Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [3]Department of Gastroenterological Surgery, Hyogo College of Medicine, Nishinomiya, Japan; [4]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Hyogo College of Medicine, Nishinomiya, Japan; [5]Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan and [6]Department of Gastroenterology and Hepatology, Osaka Red Cross Hospital, Osaka, Japan

*To whom correspondence should be addressed. Tel: +81-75-751-4319; Fax: +81-75-751-4303; Email: atsushit@kuhp.kyoto-u.ac.jp

Abstract

Hepatocellular carcinoma (HCC) developing after hepatitis C virus (HCV) eradication is a serious clinical concern. However, molecular basis for the hepatocarcinogenesis after sustained virologic response (SVR) remains unclear. In this study, we aimed to unveil the transcriptomic profile of post-SVR liver tissues and explore the molecules associated with post-SVR carcinogenesis. We analysed 90 RNA sequencing datasets, consisting of non-cancerous liver tissues including 20 post-SVR, 40 HCV-positive and 7 normal livers, along with Huh7 cell line specimens before and after HCV infection and eradication. Comparative analysis demonstrated that cell cycle- and mitochondrial function-associated pathways were altered only in HCV-positive non-cancerous liver tissues, whereas some cancer-related pathways were up-regulated in the non-cancerous liver tissues of both post-SVR and HCV-positive cases. The persistent up-regulation of carcinogenesis-associated gene clusters after viral clearance was reconfirmed through *in vitro* experiments, of which, *CYR61*, associated with liver fibrosis and carcinogenesis in several cancer types, was the top enriched gene and co-expressed with cell proliferation-associated gene modules. To evaluate whether this molecule could be a predictor of hepatocarcinogenesis after cure of HCV infection, we also examined 127 sera from independent HCV-positive cohorts treated with direct-acting antivirals (DAAs), including 60 post-SVR-HCC patients, and found that the elevated serum Cyr61 was significantly associated with early carcinogenesis after receiving DAA therapy. In conclusion, some oncogenic transcriptomic profiles are sustained in liver tissues after HCV eradication, which might be a molecular basis for the liver cancer development even after viral clearance. Among them, up-regulated *CYR61* could be a possible biomarker for post-SVR-HCC.

Introduction

Hepatocellular carcinoma (HCC) is one of the most lethal malignancies, and one of its major causes is hepatitis C virus (HCV) infection (1–3). The recent development of anti-HCV treatment, including interferon (IFN) and direct-acting antivirals (DAAs) have resulted in the achievement of sustained virologic response (SVR) in most HCV-infected patients (4,5). However, patients who have achieved HCV eradication have been reported to

Received: September 26, 2020; **Revised:** January 12, 2021; **Accepted:** February 16, 2021

Abbreviations

DAAs	direct-acting antivirals
DAVID	Database for Annotation, Visualization and Integrated Discovery
DEG	differentially expressed gene
ES	enrichment score
GO	gene ontology
GSEA	Gene Set Enrichment Analysis
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
IFN	interferon
NES	normalized enrichment score
SVR	sustained virologic response
WGCNA	weighted gene correlation network analysis

have a subsequent risk of HCC development (6–8). The average duration for HCC development after SVR (herein, post-SVR-HCC) is approximately 5 years, ranging from less than 1 year to more than 10 years (9), suggesting that both patients who achieved SVR using only the oral DAA treatment (DAA-SVR) and those using the IFN-based therapy (IFN-SVR) are at risk for HCC development for a long time after SVR. Because of the increasing number of HCV patients receiving DAA therapy, post-SVR-HCC is becoming one of the most pressing clinical issues about liver diseases worldwide. However, it remains unclear why HCC develops after the elimination of HCV as a possible carcinogen (6).

Currently, the potential risk factors for post-SVR-HCC have been investigated extensively (10,11). For example, the clinical factors associated with liver fibrosis, such as elevated serum M2BPGi value and FIB-4 index, are associated with carcinogenesis after SVR (6,10,12). Furthermore, a genome-wide association study (GWAS) on post-SVR-HCC patients revealed that a single-nucleotide polymorphism (SNP) on the *TLL1* gene involved in the liver fibrogenesis pathway could be associated with the development of post-SVR-HCC (13). To further understand the development of post-SVR-HCC and establish its clinical biomarkers, molecular profiling studies on liver tissues associated with risk factors for post-SVR-HCC are warranted. However, there are only a few studies describing the transcriptomic profile of post-SVR-HCC patients (14,15), and none of them have yet been applied to clinical practice.

Considering the carcinogenesis in inflamed tissues, the significance of the oncogenic potential of non-tumour tissues has been recently examined in several organs. For example, the clonal expansion in non-tumour cells harbouring certain cancer-related somatic mutations has been determined in the oesophagus and colon (16,17). These findings are connected with the so-called field cancerization. Similarly, many liver cancers develop from inflamed liver tissues, such as chronic hepatitis or cirrhosis. We have previously reported that non-tumour cirrhotic liver tissues accumulate genetic aberrations, including cancer-related somatic mutations (18). Although not all of these somatic mutations are associated with hepatocarcinogenesis, genetic analyses have indicated their oncogenic potential. Importantly, these oncogenic aberrations in non-tumoural tissues may be potential therapeutic targets or predictive biomarkers for cancer prognosis or development (19,20).

Therefore, in this study, we aimed to unveil the oncogenic potential after SVR achievement through the transcriptomic profiling of post-SVR non-tumour liver tissues and the HCV-infected cell line experiment using RNA sequencing (RNA-seq) (21). Furthermore, using the integrated data of clinical and *in vitro* samples, we explored the molecules whose abnormal expression levels persist after HCV clearance and validated their potential as predictive biomarkers for post-SVR-HCC using serum samples from the patients receiving DAA therapies.

Materials and methods

Patients and sample collection

Of the 247 patients who underwent hepatectomy between January 2009 and June 2019 at the Kyoto University Hospital, 25 were diagnosed with post-SVR-HCC by contrast-enhanced computed tomography or magnetic resonance imaging (Gd-EOB-DTPA MRI) according to the strict clinical criteria. Samples of their non-tumourous liver tissues were obtained. Two patients whose liver samples were not available for RNA extraction and three patients whose RNA samples were not of high quality for RNA-seq were excluded. Consequently, 20 non-tumour liver samples, including 11 from IFN-SVR patients and 9 from DAA-SVR patients were subjected to RNA-seq. For analysing transcriptional heterogeneity, 14 liver tissues were randomly collected from the non-tumoural cirrhotic livers of 4 HCV-positive patients (5 each from 2 patients and 2 each from the others) and subjected to RNA-seq (Supplementary Figure S1 and Supplementary Table S1, available at *Carcinogenesis* Online). As control, 7 normal liver tissues from healthy donors of living-donor liver transplantation resected at Kyoto University Hospital were analysed.

For serum analysis, a total of 127 serum samples were collected from HCV-positive patients treated with DAA regimens. The training cohort consisted of patients treated at Kyoto University Hospital ($N = 56$) including 20 post-DAA-SVR-HCC patients, whereas the validation cohort consisted of those treated at Osaka Red Cross Hospital ($N = 71$) including 40 post-DAA-SVR-HCC patients. Details are described in the Supplementary Materials and methods, available at *Carcinogenesis* Online.

This research conformed to the provisions of the Declaration of Helsinki, and the study protocol was approved by the ethics committee of the Kyoto University and Osaka Red Cross Hospital. Written informed consent was obtained from each patient, while we also applied opt-out method to obtain consent if applicable.

Total transcriptome analysis

RNA-seq was conducted with the HiSeq4000/NovaSeq 6000 platform (Illumina). Sequence data analyses including alignment and raw read count were performed using Genomon RNA pipeline (22,23), and differentially expressed genes (DEGs) were extracted by edgeR (24). Details are described in the Supplementary Materials and methods, available at *Carcinogenesis* Online.

Pathway analysis

Pathway analysis was conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources 6.8 software (25) with default settings. Gene Set Enrichment Analysis (GSEA) was performed using a public software obtained from the Broad Institute to compare the gene expression profiles of each case (26). Significantly enriched gene sets were clustered and visualized using Cytoscape.

Reanalysis of RNA-seq data in the public database

We selected 40 datasets of Japanese patients with HCV-positive HCC (HCV-HCC) registered in the International Cancer Genome Consortium (ICGC) using the propensity score matching method and included the following clinical factors: patients' age, sex, tumour stage and degree of fibrosis (METAVIR score). Using the R software, 1:2 assignment was conducted. The RNA-seq datasets of 40 non-tumour tissues from HCV-HCC patients were downloaded from the ICGC Data Portal (https://dcc.icgc.org/). FASTQ files of each dataset were processed and analysed using the same protocol as our dataset.

Cell line experiments

Huh7.5.1.-8 (referred to as Huh7) cells were kindly gifted by Dr Fukasawa (27) in November 2017 and authenticated via whole genome sequencing

by the report (28). We transfected *in vitro* transcribed full-length JFH-1 RNA into naive Huh7 cells, a human hepatoma cell line (29). The culture medium was collected from the cells, inoculated in Huh7 cells, and the viral RNA was quantified using RT–PCR with real-time detection (30). We also confirmed by immunofluorescence analysis that almost all JFH-1-infected cells were positive for core proteins before using infected cells for treatment (31). Then, the HCV-infected cells were treated with 2 μM of sofosbuvir for 30 days and were confirmed the complete clearance of HCV-RNA from the cells. Total RNA was extracted as previously described and subjected to library construction for RNA-seq. The experiments were performed in triplicates.

Statistical analysis

Statistical analysis was performed using R v4.0.1. Categorical values were analysed using the Fisher's exact test, whereas continuous values were analysed using the Mann–Whitney *U*-test. Survival analysis was performed using the Kaplan–Meier method and log-rank test. Multivariate analysis was conducted using Cox proportional hazard model. Differences were considered statistically significant when $P < 0.05$. DEGs have *P*-values <0.05 and false discovery rate (FDR) <0.05.

Histopathological examination, RNA extraction, RNA-seq, DAVID software setting, weighted gene correlation network analysis (WGCNA), TCGA and Genotype-Tissue Expression (GTEx) data analysis, serum analysis of Cyr61 protein, enzyme-linked immunosorbent assay and survival analysis are described in the Supplementary Materials and methods, available at *Carcinogenesis* Online.

Results

Clinicopathological characteristics of study patients

Our workflow from sample collection to RNA-seq is summarized in Figure 1. The clinical characteristics of the 20 post-SVR-HCC patients (15 men, 5 women) are summarized in Supplementary Table S2, available at *Carcinogenesis* Online. Briefly, their mean age at HCC diagnosis was 70.8 years (range: 59–80), and the median duration from the SVR achievement to HCC diagnosis was 30.8 months (range: 1–180). Eleven achieved SVR using IFN-based therapy before 2014, whereas 9 developed HCC after oral DAA therapy. The DAA regimen included asunaprevir plus daclatasvir in 4 cases; sofosbuvir, ledipasvir, 2; sofosbuvir, ribavirin, 1; glecaprevir, pibrentasvir in 1; and elbasvir, grazoprevir, 1. Almost all (19/20) patients were diagnosed at TNM stage 1 or 2. In addition, 14 tumours were moderately differentiated; 2, well-differentiated; and 4, poorly differentiated. Among non-tumourous liver tissues, the number of cases with METAVIR fibrosis grade F0/1/2/3/4 was 0/2/3/6/9, respectively, whereas the inflammation activity grade A0/1/2/3 was 4/13/2/1, respectively. Propensity score matching of the 40 HCV-positive HCC Japanese patients are summarized in Supplementary Table S3, available at *Carcinogenesis* Online.

To validate the transcriptomic alterations in DAA treatment, we performed RNA-seq on 9 samples from Huh7 cell lines before and after JFH-1 infection and after JFH-1 eradication by DAA (3 samples each). Overall, RNA-seq from 90 samples was included in the current study. The mapped read information of RNA-seq is shown in Supplementary Figure S2, available at *Carcinogenesis* Online.

Transcriptional aberrations accumulated in the background liver tissues of post-SVR-HCC patients

We first examined the transcriptome profiles of non-tumour tissues from anatomically different regions of identical cirrhotic livers using 14 cirrhotic liver specimens from 4 HCV-positive patients. Although the expression levels of several genes varied, samples collected from the same liver shared similar expression profiles irrespective of their anatomic location (Figure 2A; Supplementary Figure S1, available at *Carcinogenesis* Online).

Based on these data, we conducted RNA-seq of surgically resected specimens obtained by random sampling of non-tumour liver tissues from SVR-HCC cases. We identified 5340 and 5546 DEGs between post-SVR versus normal livers and post-SVR versus HCV-positive liver tissues, respectively, using edgeR (FDR < 0.05 and family-wise error rate *P*-value < 0.01) (Figure 2B; Supplementary Figure S3 and Supplementary Table S4, available at *Carcinogenesis* Online). Principal component analysis demonstrated that the transcriptional status of post-SVR livers is intermediate between normal and HCV-positive livers with relatively wide distribution, whereas HCV-positive livers and the normal controls were separately plotted (Figure 2C). These data suggest that the total transcriptional profiles of post-SVR livers were slightly heterogeneous owing to their clinical background.

Further, we conducted GSEA using 15 960 gene sets deposited in MSigDB (26) and found that the most enriched gene sets in post-SVR livers were associated with immune response, such as GO_ADAPTIVE_IMMUNE_RESPONSE [normalized enrichment score (NES): 2.13 and FDR: <0.0001] and ACTIVATION_OF_IMMUNE_RESPONSE (NES: 2.07 and FDR: <0.000055) and include *FAS*, *JAK2* and *NFKB* (Figure 2D; Supplementary Figure S4 and Supplementary Tables S5 and S6, available at *Carcinogenesis* Online). This supports the consistent inflammation in post-SVR liver tissues caused by the previous HCV infection. Using HALLMARK gene sets, GSEA revealed that significantly enriched gene sets in post-SVR-HCC patients were not only inflammation-related but also cancer-related, such as IL6_JAK_STAT3_SIGNALING, MYC_TARGETS_V1 and KRAS_SIGNALING_UP (NES: 2.82, 2.58 and 2.46, respectively) (Figure 2D; Supplementary Figure S4, available at *Carcinogenesis* Online), indicating up-regulated cellular proliferation in post-SVR liver tissues.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (32) demonstrated the up-regulation of the pathways associated with cell adhesion molecules, cell cycle, inflammation and cancer (Supplementary Figure S4, available at *Carcinogenesis* Online). Network analysis using GSEA datasets based on KEGG revealed that cancer- and immunity-associated pathways formed the major clusters significantly enriched in post-SVR livers (Figure 2E). On the other hand, some metabolic pathways, such GO_ORGANIC_ACID_CATABOLIC_PROCESS (NES: –2.42 and FDR: 0.0025), were recovered in SVR cases compared with HCV-positive cases. This is in agreement with the general clinical implication (Supplementary Figure S4 and Supplementary Table S6, available at *Carcinogenesis* Online).

Oncogenic transcriptional profiles related to early carcinogenesis after SVR achievement

As one of the most important roles of HCV eradication is the suppression of HCC development for years, we examined the gene expression profiles associated with early carcinogenesis after SVR. For this purpose, we compared total transcriptome profile of non-cancerous liver tissues of post-SVR-HCC patients who developed HCC within 3 years after SVR (herein, 3y-HCC, $n = 12$) and those who achieved cancer-free survival at least 3 years (3y-noHCC, $n = 8$). GSEA using HALLMARK genes revealed that in the 3y-HCC group (Figure 2F), the liver cancer-associated molecular pathway TGF_BETA_SIGNALING (NES: 2.08, FDR: <0.0001), whose leading-edge genes contained *CTNNB1*, *ACVR1* and *TGFBR1* and the KRAS_SIGNALING_UP (NES: 1.67, FDR: 0.0067), with core enrichment of well-known cancer-related genes such as *FLT4*, *DOCK2*, *CCND2* and *MYCN* gene sets were enriched,

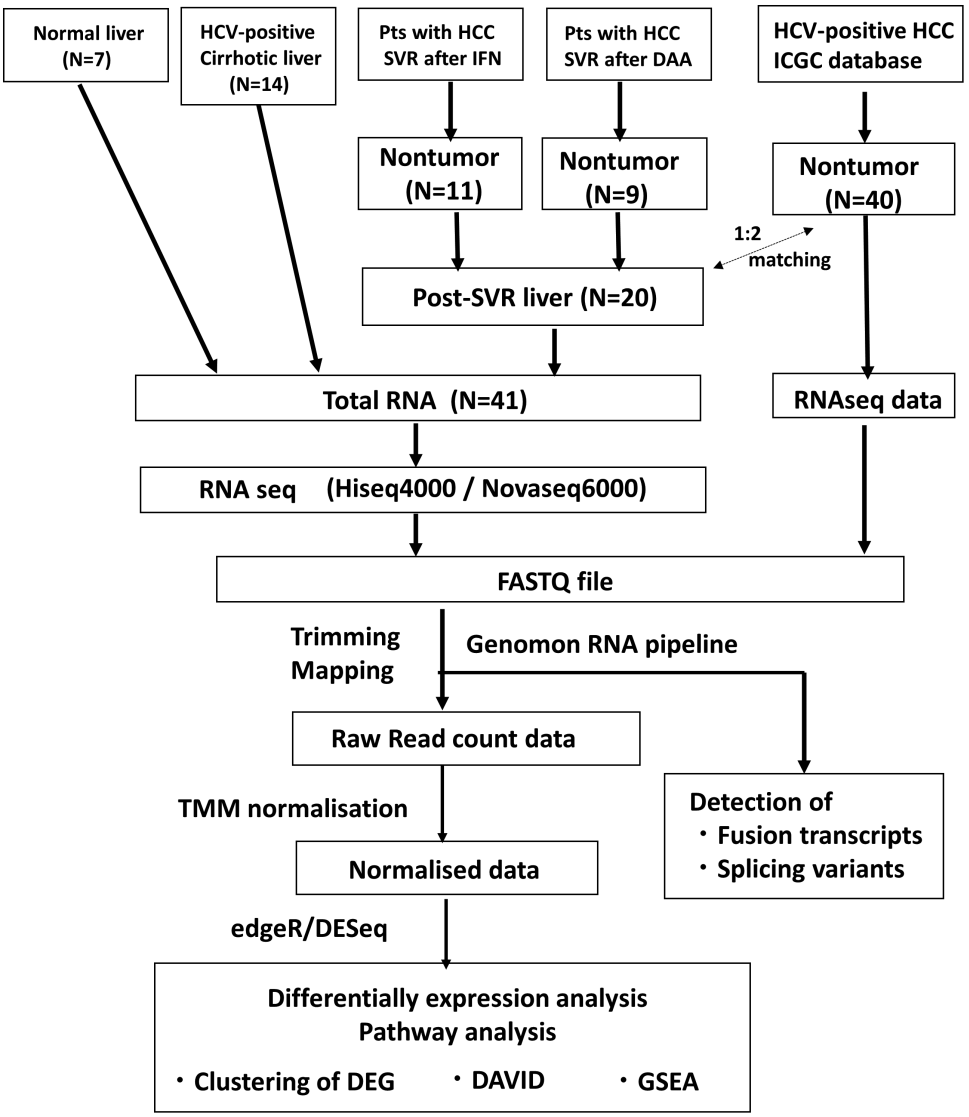

Figure 1. Workflow for clinical sample collection and transcriptomic analysis.

whereas several metabolic pathways, such as XENOBIOTIC_METABOLISM (NES: –2.22, FDR: <0.0001) with core enrichment of *APOE* and *CYP27A1*, and REACTIVE_OXYGEN_SPECIES_PATHWAY (NES:–3.39, FDR: <0.0001), were significantly down-regulated, suggesting that the 3y-HCC group are more likely to have liver dysfunction (Figure 2F). These results indicate that the non-tumour liver tissues of patients who developed HCC in a relatively early phase after SVR achievement harbour oncogenic transcriptional profiles and prolonged functional damage.

Furthermore, non-tumour livers of post-SVR patients who developed HCC within 1 year after SVR achievement (1y-HCC, $n = 5$) showed more enriched oncogenic profiles, e.g. PI3K-AKT-mTOR signalling, TGF-β signalling and other cell proliferation-associated gene sets (Figure 2G). Leading-edge analysis using the gene sets positively enriched in 1y-HCC group revealed that *NOTCH2*, *TGF-β1*, *IFNGR2*, *IL2RG* and *ITGB3* are leading-edge genes in more than three gene sets, suggesting that they are the candidates associated with early carcinogenesis after SVR.

We further examined the transcriptional profiles of liver tissues of DAA-SVR and IFN-SVR patients. The number of DEGs between these two groups was only around 400 (less than 1/10 of that between post-SVR-HCC and normal livers), suggesting fewer significantly different gene expression signatures at least on HCC development (Supplementary Table S7, available at *Carcinogenesis* Online). GSEA using MSigDB C6 (cancer-related gene sets) demonstrated that some gene sets associated with carcinogenesis, such as those whose expression levels were co-elevated with *MYC* (MYC_UP.V1_UP, NES: 2.24, FDR: <0.0001) and *CCND1* (CYCLIN_D1_UP.V1_UP, NES: 2.07, FDR: 0.00073) were enriched in DAA-SVR patients (Supplementary Table S7, available at *Carcinogenesis* Online). The results suggest that the cell cycle-associated pathway is up-regulated in DAA-SVR patients.

Differentially expressed molecular pathways in HCV-positive and post-SVR liver tissues

Although the mRNA levels of 4337 genes were up-regulated in the non-tumour liver tissues of HCV-HCC group (versus normal control), 3332 (77%) of these genes were normalized at post-SVR status (Figure 3A). DEGs among the three groups were clustered into six groups using K-means algorithm. Whereas clusters 4 and 6 were up-regulated only in HCV-positive cases, cluster 5 was up-regulated in both HCV-positive and post-SVR

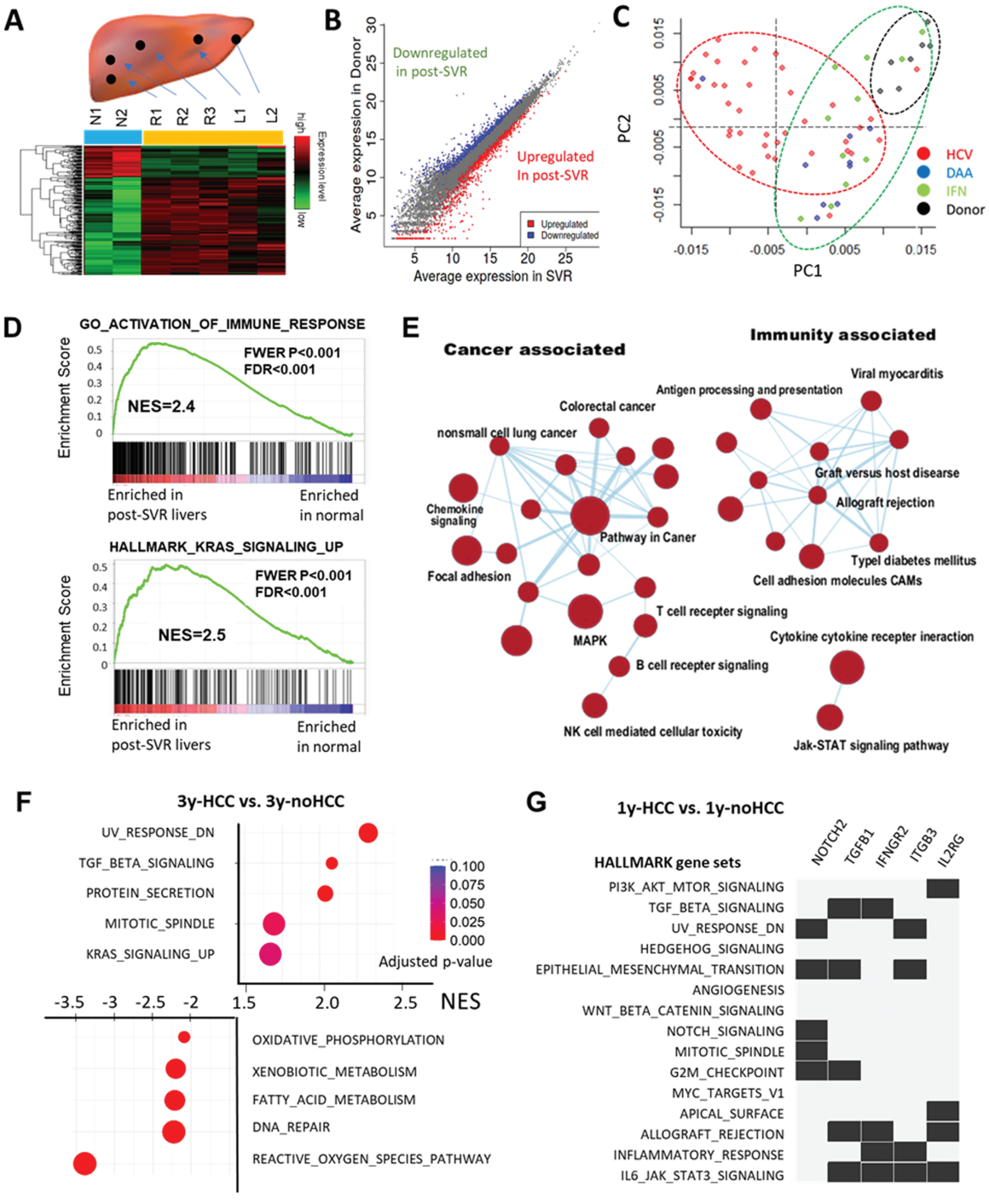

Figure 2. Transcriptomic analysis of non-tumour liver tissues of post-SVR-HCC patients. (**A**) Multiregional RNA-seq of whole explanted liver of a recipient of living-donor transplantation (LC1 case). The sampling locations are plotted in the upper schema (R1-3 from right lobe and L1-2 from left lobe). The lower heatmap represents the gene expression levels of the differentially expressed genes (DEGs) in normal livers of donors for liver transplantation (Normal1 and Normal2) and cirrhotic livers from recipients (R1-3 and L1-2). The colour key on the right of the heatmap represents the expression levels. (**B**) Scatter plots of the DEGs in post-SVR and normal livers. (**C**) Principal component analysis of non-tumour liver tissues of 40 HCV-positive, 20 post-SVR (11 IFN-SVR and 9 DAA-SVR), and normal livers. (**D**) Enrichment curve of the GO_ACTIVATION_OF_IMMUNE_RESPONSE and HALLMARK_KRAS_SIGNALING_UP, which are the representative enriched gene sets in post-SVR-HCC patients versus control. (**E**) Network of KEGG pathways enriched in post-SVR livers (FDR < 0.01) and visualized using Cytoscape, with the size of dots corresponding to the number of genes in each set. Only the representative gene sets are labelled. (**F**) GSEA of patients developing HCC within 3 years after SVR (3y-HCC) and those who did not develop HCC at least 3 years after SVR (3y-noHCC). The upper plots show the top 5 MSigDB HALLMARKS gene sets enriched in 3y-HCC cases, whereas the lower plots are those enriched in 3y-noHCC cases. The x-axis corresponds to NES. (**G**) GSEA of post-SVR patients developing HCC within 1 year after SVR and those who did not develop HCC at least 1 year after SVR. The results of the leading-edge analysis of significantly enriched gene sets in those developing HCCs within 1 year are shown. The rows show the name of significantly enriched gene sets, whereas the column shows the leading-edge genes common to more than three gene sets. In each column, the cells are in black when the gene is included in the leading edge of the corresponding gene sets.

patients (Figure 3B). Pathway and gene ontology (GO) analyses using DAVID Bioinformatics Resources 6.8 software (25) revealed that the up-regulated genes only in HCV-positive livers were mainly annotation clusters associated with DNA damage (enrichment score [ES]: 13.19), including GO_DNA_REPAIR (Benjamini-adjusted *P*-value: 8.23E-16, FDR: 2.05E-14) and cell cycle (ES: 8.62), including GO_CELL_DIVISION (adjusted *P*-value: 0.0024, FDR: 0.0054) (Figure 3C; Supplementary Table S8, available at *Carcinogenesis* Online). On the other hand, those down-regulated only in HCV-positive livers were categorized into metabolic pathways associated with mitochondrial function (ES: 6.45), including GO_MITOCHONDRION (adjusted *P*-value: 0.007, FDR: 0.017).

Next, we focused on DEGs whose expression levels were altered in HCV-positive livers and remained as such even after SVR achievement (the majority of which are in cluster 5 in Figure 3B). There were 777 genes whose expression levels were elevated in both HCV-positive and post-SVR livers compared with normal livers. One of the most enriched includes those associated with immunity, e.g. 'IMMUNITY' in UniProt Keywords (UP_KEYWORDS) (FDR: 1.63E-36) and 'INTERFERON-GAMMA-MEDIATED SIGNALING PATHWAY' in GO_BP database (FDR: 6.16E-21), suggesting that the gene expression changes caused by HCV infection are prolonged even after HCV eradication (Figure 3D; Supplementary Table S9, available at *Carcinogenesis* Online). These gene clusters included several cancer-associated genes, such as *CXCL6* coding the chemokine CXCL6 associated with cancer cell survival and metastasis. In addition, the cancer-related gene clusters were significantly up-regulated in both HCV-positive and post-SVR liver tissues, e.g. 'cell adhesion molecules (CAMs)' in KEGG pathway (FDR: 4.91E-7), including *ICAM1* coding cell adhesion factor ICAM1 and 'POSITIVE REGULATION OF ERK1 AND ERK2 CASCADE' in GO_BP database (FDR: 5.9E-5). These results suggest that some oncogenic expression profiles acquired during chronic HCV infection might not revert even after the complete clearance of HCV, although some genes involved in the cell cycle and immunity pathways were suppressed in post-SVR versus HCV-positive conditions (Figure 3E; Supplementary Figures S5 and S10, available at *Carcinogenesis* Online).

We next explored whether any DEGs are localized at certain chromosomal regions using the PREDA algorism (33). Both the up-regulated and down-regulated genes were widely distributed over the whole genome region, and there were few local concentrations of DEGs upon comparison of HCV-positive and post-SVR livers with the control (Supplementary Figure S6, available at *Carcinogenesis* Online).

Structural aberrations of mRNA, such as fusion transcripts and splicing variants, were also examined (Figure 1). Splicing variants, such as intron retention of each mRNA, of some cancer-related genes were detected in the non-tumour tissues of DAA-SVR and IFN-SVR patients. Compared with the control, however, there were no statistically significant splicing variants frequently detected in post-SVR livers. Similarly, although a few fusion transcripts were detected in post-SVR liver tissues, we could not detect commonly existing fusions.

Hepatocyte-specific transcriptional alteration during HCV infection and eradication

Transcriptome analysis on the clinical specimens through bulk sampling (so-called bulk RNA-seq) elucidates the expression profiles of whole liver tissues, which include not only hepatocytes but also non-hepatocytes such as fibroblasts and infiltrating lymphocytes. As HCV infection induces the infiltration of immune cells into liver tissues, bulk RNA-seq is especially influenced by the degree of immune cell infiltration. To determine the transcriptional changes before and after HCV infection and after HCV eradication in hepatocytes, we conducted *in vitro* experiments using the Huh7 cell line.

HCV JFH-1 was first infected into Huh7, and after the HCV-RNA level has been confirmed to be elevated, JFH-1 was completely eradicated using the sofosbuvir treatment (Supplementary Figure S7, available at *Carcinogenesis* Online). Total RNA was extracted from cells before and after JFH-1 infection and after viral eradication, and RNA-seq was performed (n = 3 for each) (Figure 4A). Similar to the analysis of clinical liver specimens, the DEGs in each group were identified (Figure 4B and C; Supplementary Figure S7, available at *Carcinogenesis* Online), and the K-means clustering algorithm classified them into four representative gene clusters according to their expression patterns (Figure 4D and E). Similar to the clinical data, the expression level of the genes elevated after JFH-1 infection (such as clusters A and C) returned to normal after viral eradication. These included inflammation-associated genes (e.g. HALLMARK INTERFERON ALPHA RESPONSE, adjusted *P*: 6.3E-13), despite the absence of immune cell infiltration (Figure 4F), suggesting that the up-regulation of inflammatory gene expression levels in the clinical samples is caused by not only the infiltration of immune cells but also the alteration in the expression profile of hepatocytes. In contrast, cluster A, whose expression level was down-regulated only in HCV-infected cells, was associated with metabolic pathways (HALLMARK XENOBIOTIC METABOLISM, adjusted *P*: 1.3E-3). Liver-specific genes were also down-regulated in this cluster (Figure 4F).

On the other hand, there were genes whose expression levels increased after JFH-1 infection and remained even after JFH-1 eradication, such as those in cluster B (Figure 4E), which was also observed in clinical samples (Figure 3E). Interestingly, pathway analyses revealed that this cluster was significantly associated with cell mortality, growth factor activity and some oncogenic signatures, like the Hippo signalling pathway (Figure 4F).

CYR61 and co-expressed oncogenic genes are overexpressed after HCV eradication

Next, we compared gene cluster B in Figure 4E with those having the same expression patterns in HCV-positive and post-SVR liver tissues. Consequently, we identified 15 genes whose expression levels were elevated during HCV infection and after HCV eradication in the clinical samples and *in vitro* experiments (Figure 5A; Supplementary Table S11, available at *Carcinogenesis* Online). Among these, *CYR61*, a member of the cellular communication network (CCN) family, has been reported to be associated with carcinogenesis in many types of cancers (34–36), whereas *CXCL6* is one of the major inflammation-associated genes. These 15 genes also included *LOXL4*, which has been identified as one of the genes whose expression is elevated in HCV-infected and post-SVR status, as reported in a previous study (15). In contrast, we found eight genes whose expression levels were suppressed during HCV infection and after HCV eradication in the clinical samples and *in vitro* experiments, none of which have been reported to be associated with cancer development.

We focused on *CYR61* as it is the top-ranked DEG with such an expression profile (FDR < 1×10^{-54}). We originally constructed a gene set associated with *CYR61* based on the previously reported gene expression database Gene Expression Omnibus

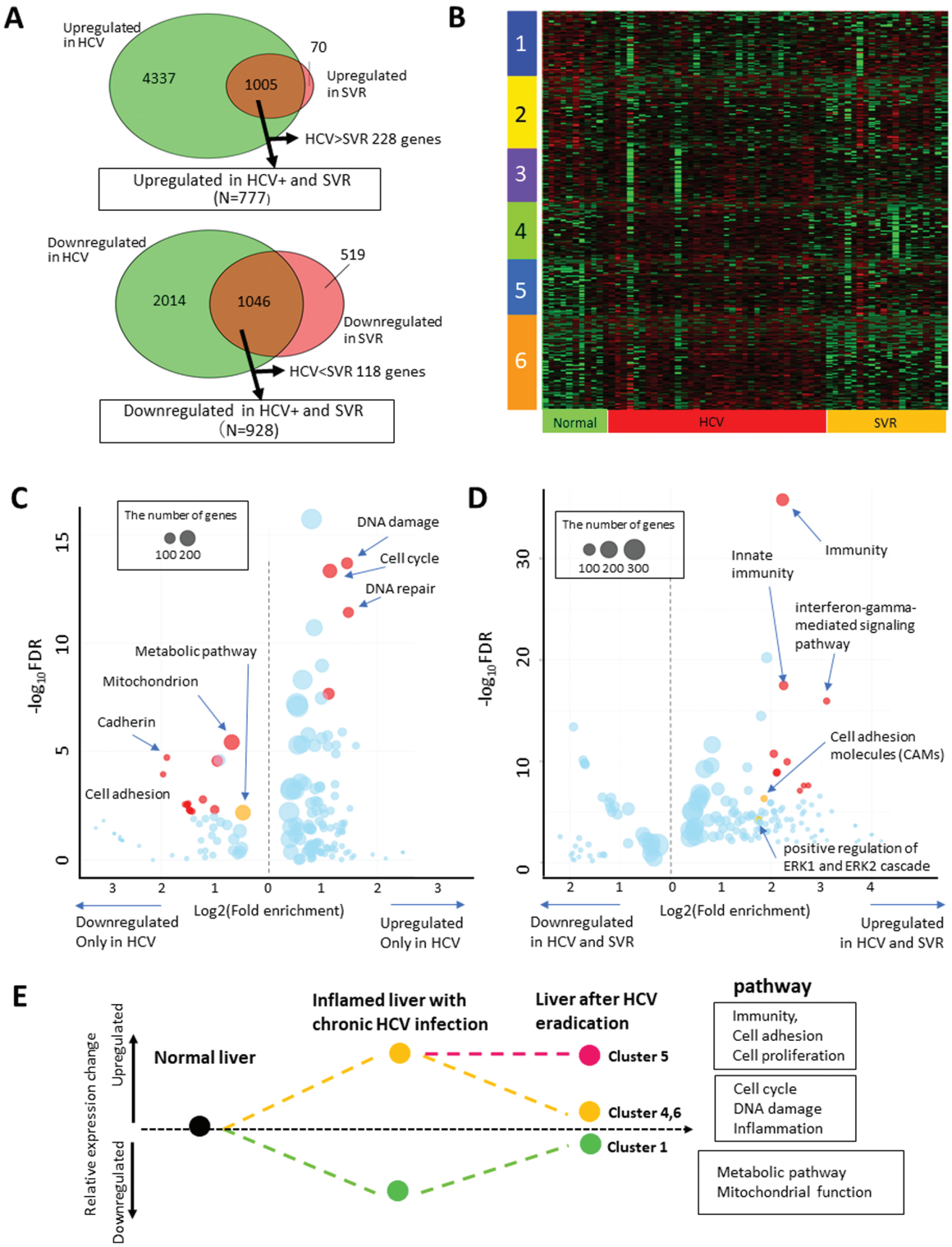

Figure 3. Functional analysis of the differentially expressed genes (DEGs) in post-SVR-HCC, HCV-HCC and control. (**A**) Venn diagrams of DEGs which were elevated (upper diagram) and decreased (lower diagram) in HCV-positive and post-SVR patients versus normal liver tissues. (**B**) Heatmap of the expression levels of DEGs. Clustering was conducted using the K-means algorism, and six gene clusters were detected. (**C**) Dot plots showing the up- or down-regulated pathways only in HCV-positive livers versus normal livers and non-tumour tissues of HCV-positive and post-SVR patients using DAVID. Genes categorized into clusters 4 or 6 in B were included. Each dot represents significantly enriched pathway, with its size corresponding to the number of genes included in each cluster. Only the pathways with FDR < 0.01 are plotted. Based on the functional cluster analysis, the major pathways of the top enriched annotation clusters are in red, and pathways of interest are in orange. (**D**) Pathways that are up- or down-regulated in the non-tumour liver tissues of both HCV-HCC and post-SVR-HCC patients versus control (included in cluster 5 in B). The x-axis shows $\log_2$ (fold enrichment) of each pathway, and y-axis represents $-\log_{10}$ (FDR). (**E**) The scheme of the representative gene expression patterns.

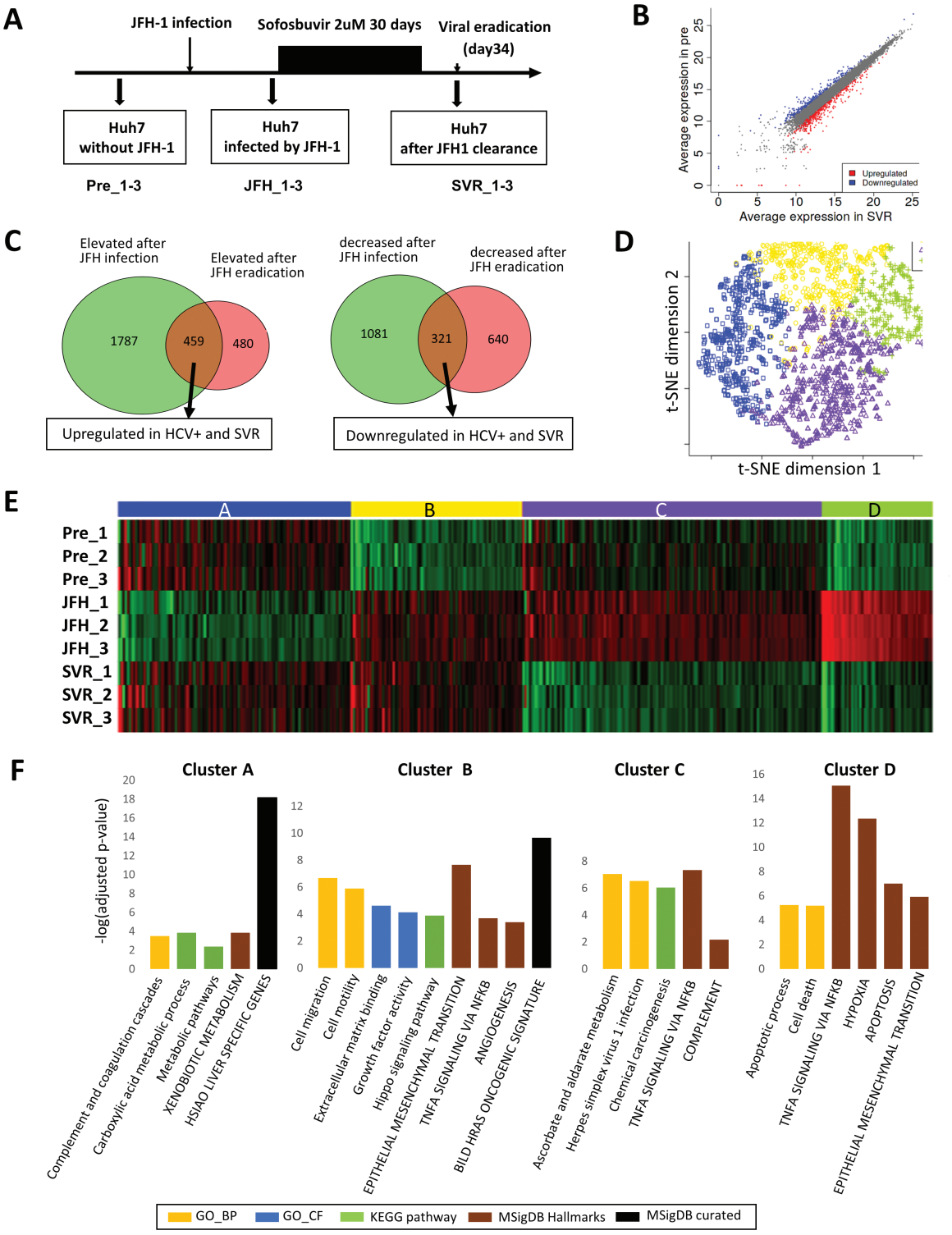

Figure 4. *In vitro* validation of differentially expressed genes (DEGs). (**A**) Workflow of the cell line experiment. Total RNA was extracted from Huh7 before and after JFH-1 infection and after SVR using DAA therapy ($N = 3$ for each). (**B**) Scatter plot of DEGs in pre-infected (blue) and post-SVR (red) cells. (**C**) Venn diagrams of DEGs that were elevated (left) and decreased (right) in HCV-infected status and post-SVR status compared with pre-infection status. (**D**) t-Distributed Stochastic Neighbor Embedding plot of the top significant 2000 DEGs divided into four clusters. (**E**) Heatmap of the top 2000 DEGs among three groups. Clustering was conducted using the K-means algorism, and DEGs were categorized into four clusters. (**F**) Major gene sets and pathways are enriched in each gene cluster. Gene Ontology Biological Process (GO_BP) and Cellular Function (BP_CF), KEGG pathway, MSigDB Hallmarks and curated gene sets are used in this analysis. *y*-axis shows the $\log_{10}$ (adjusted *P*-value).

(GEO). The interaction network of *CYR61* included genes associated with cell adhesion, such as those from the integrin family (*ITGR5* and *ITGR6*), cell proliferation (*IGFBP2*), extracellular matrix (*MMP9*) and inflammation (*STAT3* and *IL6*) (Figure 5B).

From the GSEA using this gene set, we found that the *CYR61*-associated gene set was positively enriched in both HCV-positive and post-SVR conditions compared with the controls in both clinical samples and *in vitro* experiments (NES: 1.71 and

2.07 in the clinical samples, respectively) (Figure 5C). The expression levels of the leading-edge genes are represented in the heatmap (Figure 5D). While the expression levels of these genes were generally suppressed in the control, most HCV-positive liver tissues harboured elevated expression levels. Importantly, many of these genes were highly expressed, together with the elevated level of *CYR61* mRNA, even in the post-SVR liver tissues, thus indicating that the molecular pathways associated with *CYR61* genes, including cellular proliferation, remain up-regulated even after SVR achievement, although the degree of their activation might be suppressed compared with HCV-positive status.

Furthermore, we conducted a WGCNA (37) of the top 2000 DEGs in the *in vitro* experiments and identified 10 modules (Figure 5E). The magenta-coloured module 9 was constructed with the genes co-expressed with *CYR61* (Figure 5F). These genes were up-regulated after HCV infection and remained so even after HCV eradication. KEGG pathway and GO analyses demonstrated that this module was significantly associated with some oncogenic gene sets, such as those up-regulated by the activation of the epidermal growth factor receptor (EGFR) pathway (*P*: 0.00001), and cell proliferation (e.g. 'POSITIVE REGULATION OF CELLULAR PROCESS', *P*: 0.0001) (Figure 5G). These data suggest that *CYR61* and its co-expressed genes are associated with prolonged oncogenic potential after SVR achievement.

To determine whether up-regulation of *CYR61* mRNA level is an HCV-HCC specific event or an HCV event, we examined the *CYR61* mRNA level in liver tumour and adjacent liver tissues using RNA-seq data deposited in TCGA database ($N = 400$) and compared with that in normal livers reported in GTEx database ($N = 268$). The datasets demonstrated that *CYR61* mRNA levels in HCCs were elevated compared with those in normal livers but were significantly decreased compared with those in adjacent liver tissues (Supplementary Figure S8, available at *Carcinogenesis* Online), indicating that elevation of serum Cyr61 protein is not induced by HCC but HCV infection itself.

Serum Cyr61 as a candidate biomarker for early hepatocarcinogenesis after SVR achievement

We hypothesized that the elevated expression level of *CYR61*, even after viral eradication, in the liver tissues can be associated with carcinogenesis after SVR. Therefore, we examined the potential of serum Cyr61as biomarker for hepatocarcinogenesis after SVR achievement in the clinical setting. As a training cohort, we collected 56 serum samples from 20 patients who developed HCC within 5 years after DAA treatment (SVR-HCC group) and 36 patients who did not develop HCC at least 5 years after DAA treatment (control group) (Figure 6A). The SVR-HCC group included 3 cases analysed using RNA-seq and 17 cases from the other cohort whose tumours were treated by radiofrequency ablation in our hospital. Patient characteristics are described in Figure 6B.

We found that serum Cyr61 level was positively correlated with the FIB4-index but negatively correlated with the platelet count in the control group (Figure 6C; Supplementary Figure S9, available at *Carcinogenesis* Online). Further, the duration from the end of treatment (EOT) to HCC development was negatively correlated with serum Cyr61 level at SVR achievement (Figure 6D). Ten of 20 patients who developed HCC within 1 year after EOT had significantly higher serum Cyr61 levels at the time of HCV eradication than those who developed HCC more than 1 year after EOT (Figure 6E). Finally, survival analysis of 56 patients using Kaplan–Meier estimation revealed that patients with serum Cyr61 >350 ng/μL at SVR had significantly shorter cancer-free survival time than those with Cyr61 <350 ng/μL (Figure 6F).

Cancer-free survival time was further analysed in an independent validation cohort ($N = 71$) including 40 SVR-HCC patients (see Supplementary Materials and methods and Supplementary Table S12, available at *Carcinogenesis* Online). We confirmed that higher Cyr61 protein level was significantly correlated with poor cancer-free survival after DAA treatment (Figure 6G). Multivariate analysis with other liver disease-associated parameters demonstrated that higher serum Cyr61 level was independently associated with shorter cancer-free survival after SVR in both training and validation cohorts (Supplementary Tables S13 and S14, available at *Carcinogenesis* Online), indicating that a high level of serum Cyr61 at HCV eradication might be a potential biomarker to predict early carcinogenesis after SVR achievement.

Discussion

Through comparative analyses of the total transcriptome data of HCV-infected livers and post-SVR liver specimens, we found the persistent oncogenic expression patterns after HCV eradication. In the liver tissues with chronic HCV infection, the genes associated with mitochondrial function were significantly down-regulated, whereas those with DNA damage, cell cycle and immunity were significantly up-regulated. Pathways associated with cellular proliferation were also activated in the non-tumour tissues of HCV-HCC cases, suggesting the oncogenic potential of HCV-infected liver tissues. As expected, most of these changes in expression levels were normalized or reduced after the complete eradication of HCV. These gene expression patterns agree with the well-known dynamics of clinical data before and after HCV eradication, as represented by the improvement in liver function and the normalization of serum alanine aminotransferase (ALT) levels after anti-HCV therapies (38). On the other hand, we found that cell cycle- and cell proliferation-associated pathways remained significantly up-regulated even after the eradication of HCV. This was also the case with immunity-related genes, although they were partially normalized compared with HCV-positive cases.

To exclude the influence of immune cell infiltration on the transcriptional profile of the clinical specimens, we further examined the transcriptional changes in an HCV-infected cell line. Under the condition without the influences of non-hepatocytes such as lymphocytes, expression levels of inflammation-related genes were up-regulated in Huh7 after HCV infection and normalized after HCV clearance. Interestingly, some oncogenic associated pathways were up-regulated after HCV infection and even after HCV clearance. These findings in cell line experiment agree with the sequencing data on our clinical samples, suggesting that, once infected with HCV, some expression profiles cannot be completely erased from hepatocytes even after viral clearance, leaving some oncogenic expression signature on liver tissues. It is unclear why the elevated expression levels of some genes are sustained even after HCV clearance. Recent studies have demonstrated that some epigenetic changes caused by HCV infection can be maintained after HCV eradication (15,39). Considering these reports, some epigenetic alterations, including histone acetylation and DNA methylation on cancer-related genes, might remain in chronic hepatitis tissues owing to active HCV infection, resulting in the prolonged elevation of specific gene expressions even after HCV clearance. To verify this hypothesis, integrated analyses using transcriptomic and epigenomic data should be further conducted.

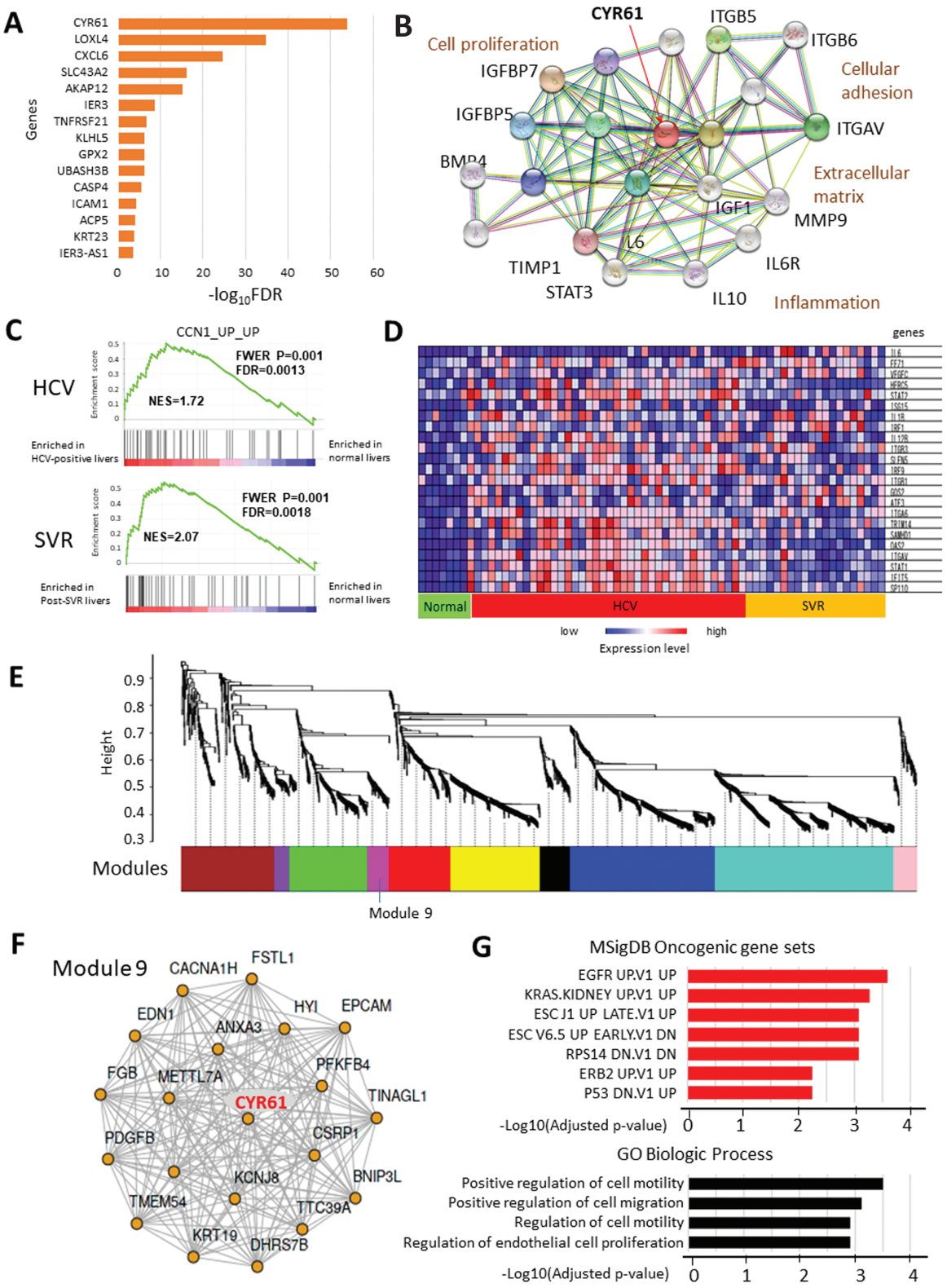

Figure 5. The elevated expression level of the *CYR61*-associated gene cluster is prolonged after SVR achievement. (**A**) Genes with significantly elevated expression levels in HCV-positive and post-SVR specimens in both *in vitro* experiment and clinical cases. (**B**). Molecular network of genes associated with the *CYR61* gene generated using STRING. (**C**) GSEA plot based on the originally generated gene set CCN1_UP_UP, including the genes whose expression levels, have been reported to be elevated when *CYR61* is highly expressed (upper: HCV-positive livers versus normal; lower: post-SVR livers versus normal). (**D**). Heatmap of expression levels of leading-edge subset genes extracted using GSEA based on the *CYR61*-associated gene set. (**E**) Gene dendrogram of weighted gene correlation network analysis (WGCNA) of the top 2000 DEGs in the cell line experiment. A total of 10 gene modules were determined. (**F**) Top 20 genes of module 9 including *CYR61* (magenta). (**G**) Top gene sets using MSigDB and GO biologic process enriched in module 9 with the x-axis showing genes $-\log_{10}$ (adjusted *P*-value).

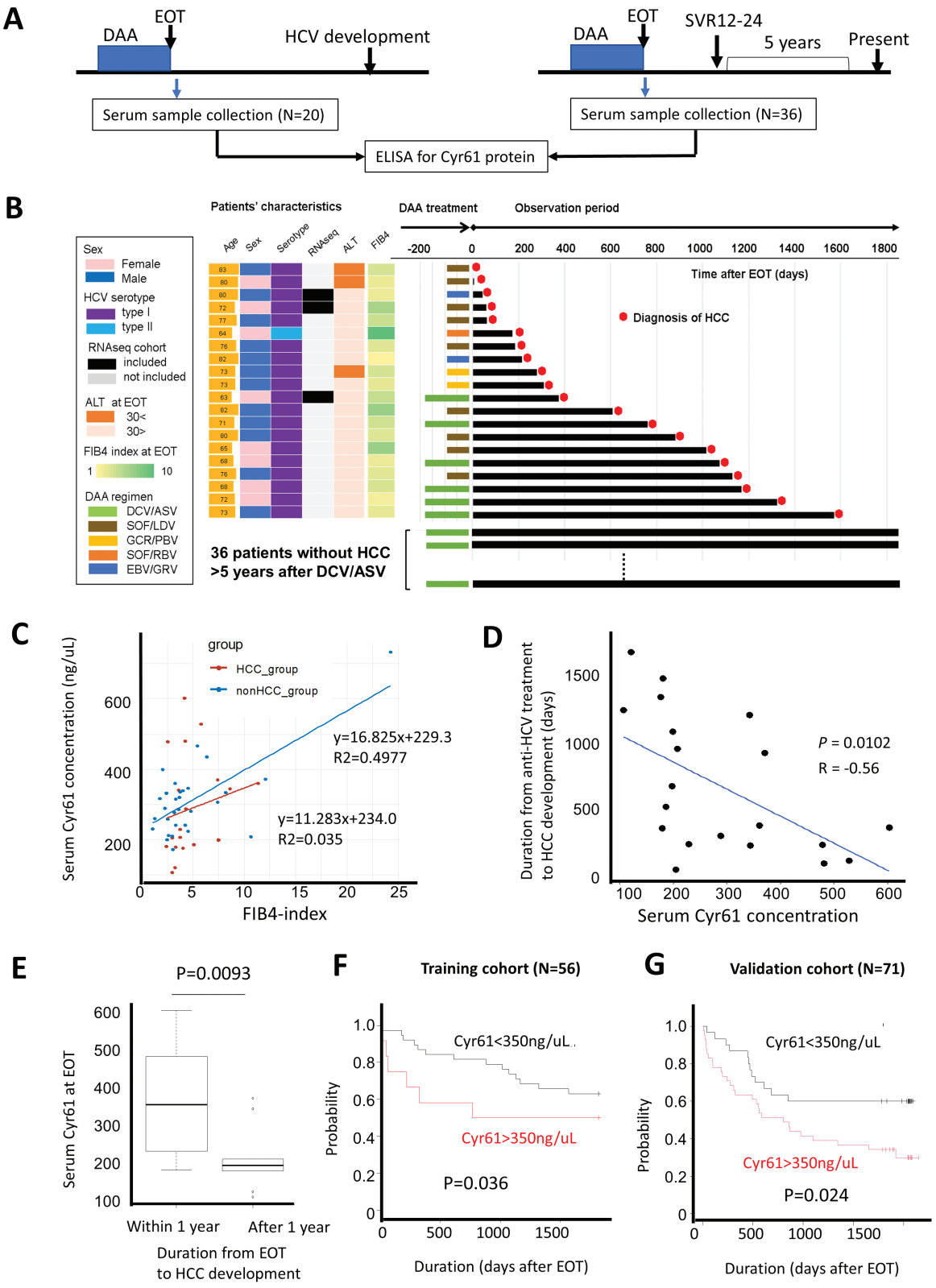

Figure 6. Validation study of Cyr61 protein as a potential predictive biomarker of post-SVR-HCC using human serum. (**A**) Workflow of sample collection. (**B**) Patient characteristics with colour keys at the left. Right bar plots show the durations of DAA treatment, observation after end of treatment (EOT), and until HCC diagnosis. The observation period of 36 patients without HCC was censored at 5 years after the EOT. (**C**) Dot plot of serum Cyr61 level and FIB4 index at HCV clearance. (**D**) Relationship between serum Cyr61 level and duration from the end of anti-HCV treatment to HCC development. (**E**) Boxplots of the serum Cyr61 concentration at EOT according to the timing of HCC development (within 1 year after SVR or not). (**F** and **G**) Survival analysis of the training cohort (F) and the validation cohort (G) using Kaplan-Meier curves. Black and red lines represent those who had serum Cyr61 level <350 ng/μL and those with more than 350 ng/μL, respectively. The x-axis represents the days after the completion of DAA therapy, and the y-axis shows the cancer-free survival rate.

Focusing on the similar expression dynamics between clinical samples and HCV-infected cell lines, we confirmed that the expression of *CYR61*, the top-ranked DEG, was elevated during chronic HCV infection and sustained even after HCV eradication. *CYR61*, also known as *CCN1*, is a member of the CCN family and has been reported to be related to carcinogenesis or metastasis in many organs, including colon, breast and ovary (34–36). Although the association between *CYR61* overexpression and liver carcinogenesis remains controversial, *CYR61* is known to be expressed predominantly in hepatocytes and activates stellate cells, which can accelerate liver fibrosis (40,41). Further, *CYR61* is reported to be regulated by the Wnt/β-catenin signalling and can promote hepatocarcinogenesis (42). In addition, *CTGF*, also known as *CCN2*, which has been reported to activate hepatocarcinogenesis (43), can interact with *CYR61* (44). Interestingly, Hamdane *et al*. described that the elevated expression of *SPHK1* remains increased after HCV eradication and demonstrated that high expression of *SPHK1* is significantly associated with HCC risk after SVR (15). Several studies have demonstrated that *SPHK1* induces high expression of *CYR61* and tumour development (45). Our data demonstrated that the *CYR61* mRNA levels in liver tissues with chronic HCV infection were not normalized after HCV eradication, supporting the results of Hamdane *et al*. WGCNA analysis demonstrated that the gene modules included those whose expression patterns correlated with *CYR61* mRNA, i.e. up-regulated during HCV infection and after SVR. KEGG pathway analysis revealed that the EGFR pathway or Wnt/β-catenin pathway was significantly up-regulated in this gene module, suggesting that cell proliferation is activated during HCV infection and does not completely normalize after its eradication. This indicates a persistent post-SVR oncogenic gene expression profile.

To elucidate the characteristics of patients who develop HCC, control patients who are cancer-free after SVR achievement should be comparatively analysed. There are, however, two obstacles to conducting this comparison using surgical specimens. First, it is ethically difficult to obtain liver tissues from cancer-free asymptomatic patients who achieved SVR. Second, 'cancer-free patients' are not absolutely definitive because of the relatively long duration between SVR achievement and cancer detection. Considering these obstacles, we compared the total transcriptome profiles of patients who developed HCC within 3 years after SVR (3y-HCC) with those who were cancer-free for at least 3 years after SVR (3y-noHCC). Interestingly, some cancer-related pathways, including the KEGG pathway genes up-regulated at KRAS activation, were significantly up-regulated in 3y-HCC group, whereas metabolic pathways were well maintained in the cancer-free group. Notably, patients developing HCC within 1 year after SVR achievement harboured a higher enrichment of these oncogenic gene expression signatures.

In addition, we examined serum samples to validate the malignant potential of *CYR61* up-regulation. Due to the availability of serum sample from post-SVR cancer-free patients, here we could compare the serum Cyr61 levels of post-SVR-HCC patients and those who are cancer-free for more than 5 years. We found that the elevated serum Cyr61 levels were significantly associated with the early development of HCC after SVR, and this association was validated in an independent cohort by multivariate analysis with Cox hazard model, suggesting that Cyr61 might be a potential biomarker to predict cancer-free survival after HCV eradication. As these cohorts consist of a small number of patients, further validation studies are warranted.

This study has some limitations. First, the total transcriptome analyses could not include any post-SVR patients without HCC due to the above-mentioned issues. Second, we only studied a small number of cases. However, by conducting cell line experiments and serum analysis, we believe we have sufficiently compensated for shortcomings in the bulk-sequenced data. The third limitation is the bulk sampling nature of the RNA-seq technique, which caused our inability to unveil cell type-specific gene expression profiles in detail. Single-cell RNA-seq could be useful to elucidate the expression profiles of each cell population, including tumour cells, immune cells, fibroblasts and stellate cells, in the microenvironment of the liver tissues of post-SVR patients (46)

In conclusion, some cancer-related transcriptional profiles including the up-regulation of cell proliferation pathway caused by chronic HCV infection can be sustained even after viral clearance. Elevated serum Cyr61 levels were associated with shorter duration from HCV clearance to HCC development after SVR and thus could be a potential predictive biomarker for the early occurrence of HCC after SVR. Novel analytical methods, such as single-cell analysis, on larger cohorts, could elucidate more detailed transcriptional profiling of liver tissues and establish predictive biomarkers for hepatocarcinogenesis after SVR.

Supplementary material

Supplementary data are available at *Carcinogenesis* online.

Acknowledgements

We thank Drs Tadashi Inuzuka, Tomonori Matsumoto, Tomoyuki Goto, Aya Mizuguchi, Minami Lee, Fumiyasu Nakamura, Soo Ki Kim, Nobuyuki Kakiuchi, Mari Teramura and Shigeharu Nakano for their interpretation of data and helpful advice. We also thank Drs Satoru Seo, Hideaki Okajima and Shinji Uemoto for their material support.

Funding

This work was supported by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research, KAKENHI (grant numbers 16K09357, 17H04158 and 19K08443) and Japan Agency for Medical Research and Development, AMED, Japan (grant number 18fk0210022h0502).

Conflict of Interest Statement: None declared.

Authors' contributions

Conceived and designed the experiment: H.T., A.T. and H.M.; Performed the experiments: H.T., E.I., S.A. and H.A.; Analysed the data: H.T., E.I., K.K. and Y.U.; Contributed reagents/materials/analysis tools: H.T., M.M., S.A., K.T., Y.U., K.T., E.H., H.I., H.A. and T.W.; Wrote the original manuscript: H.T. and A.T.; Critical revision of the manuscript for intellectual content: K.K., Y.E., T.S., K.T., Y.U., H.A. and TW; Supervised the project: H.M. and H.S.

Data availability

Sequence data were deposited at the Japanese Genotype-Phenotype Archive (JGA, http://trace.ddbj.nig.ac.jp/jga), which is hosted by the DDBJ, under accession numbers JGAS000234 and JGAS000134.

References

1. Bray, F. *et al.* (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.*, 68, 394–424.
2. El-Serag, H.B. (2012) Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology*, 142, 1264–1273.e1.
3. Takeda, H. *et al.* (2017) Genetic basis of hepatitis virus-associated hepatocellular carcinoma: linkage between infection, inflammation, and tumorigenesis. *J. Gastroenterol.*, 52, 26–38.
4. Maan, R. *et al.* (2016) Safety and effectiveness of direct-acting antiviral agents for treatment of patients with chronic hepatitis C virus infection and cirrhosis. *Clin. Gastroenterol. Hepatol.*, 14, 1821–1830.e6.
5. Takeda, H. *et al.* (2017) Evolution of multi-drug resistant HCV clones from pre-existing resistant-associated variants during direct-acting antiviral therapy determined by third-generation sequencing. *Sci. Rep.*, 7, 45605.
6. Ioannou, G.N. *et al.* (2019) Increased risk for hepatocellular carcinoma persists up to 10 years after HCV eradication in patients with baseline cirrhosis or high FIB-4 scores. *Gastroenterology*, 157, 1264–1278.e4.
7. Calvaruso, V. *et al*; Rete Sicilia Selezione Terapia–HCV (RESIST-HCV). (2018) Incidence of hepatocellular carcinoma in patients with HCV-associated cirrhosis treated with direct-acting antiviral agents. *Gastroenterology*, 155, 411–421.e4.
8. Ioannou, G.N. *et al.* (2018) HCV eradication induced by direct-acting antiviral agents reduces the risk of hepatocellular carcinoma. *J. Hepatol.*, *68*, 25–32.
9. Asahina, Y. *et al.* (2013) α-Fetoprotein levels after interferon therapy and risk of hepatocarcinogenesis in chronic hepatitis C. *Hepatology*, 58, 1253–1262.
10. Nagata, H. *et al*; Ochanomizu Liver Conference Study Group. (2017) Effect of interferon-based and -free therapy on early occurrence and recurrence of hepatocellular carcinoma in chronic hepatitis C. *J. Hepatol.*, 67, 933–939.
11. Waziry, R. *et al.* (2017) Hepatocellular carcinoma risk following direct-acting antiviral HCV therapy: a systematic review, meta-analyses, and meta-regression. *J. Hepatol.*, 67, 1204–1212.
12. Nagata, H. *et al*; Ochanomizu Liver Conference Study Group. (2016) Serial measurement of Wisteria floribunda agglutinin positive Mac-2-binding protein is useful for predicting liver fibrosis and the development of hepatocellular carcinoma in chronic hepatitis C patients treated with IFN-based and IFN-free therapy. *Hepatol. Int.*, 10, 956–964.
13. Matsuura, K. *et al*; Japanese Genome-Wide Association Study Group for Viral Hepatitis. (2017) Genome-wide association study identifies TLL1 variant associated with development of hepatocellular carcinoma after eradication of hepatitis C virus infection. *Gastroenterology*, 152, 1383–1394.
14. Nakagawa, S. *et al*; Precision Liver Cancer Prevention Consortium. (2016) Molecular liver cancer prevention in cirrhosis by organ transcriptome analysis and lysophosphatidic acid pathway inhibition. *Cancer Cell*, 30, 879–890.
15. Hamdane, N. *et al.* (2019) HCV-Induced epigenetic changes associated with liver cancer risk persist after sustained virologic response. *Gastroenterology*, 156, 2313–2329.e7.
16. Kakiuchi, N. *et al.* (2020) Frequent mutations that converge on the NFKBIZ pathway in ulcerative colitis. *Nature*, 577, 260–265.
17. Yokoyama, A. *et al.* (2019) Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature*, 565, 312–317.
18. Kim, S.K. *et al.* (2019) Comprehensive analysis of genetic aberrations linked to tumorigenesis in regenerative nodules of liver cirrhosis. *J. Gastroenterol*, 54, 628–640.
19. Brunner, S.F. *et al.* (2019) Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature*, 574, 538–542.
20. Zhu, M. *et al.* (2019) Somatic mutations increase hepatic clonal fitness and regeneration in chronic liver disease. *Cell*, 177, 608–621.e12.
21. Mardis, E.R. (2009) New strategies and emerging technologies for massively parallel sequencing: applications in medical research. *Genome Med.*, 1, 40.
22. Yoshida, K. *et al.* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, 478, 64–69.
23. Yoshizato, T. *et al.* (2015) Somatic mutations and clonal hematopoiesis in aplastic anemia. *N. Engl. J. Med.*, 373, 35–47.
24. Sun, J. *et al.* (2013) TCC: an R package for comparing tag count data with robust normalization strategies. *BMC Bioinformatics*, 14, 219.
25. Huang, D.W. *et al.* (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, 4, 44–57.
26. Subramanian, A. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*, 102, 15545–15550.
27. Shirasago, Y. *et al.* (2015) Isolation and characterization of an Huh.7.5.1-derived cell clone highly permissive to hepatitis C virus. *Jpn. J. Infect. Dis.*, 68, 81–88.
28. Kawamoto, M. *et al.* (2020) Identification of characteristic genomic markers in human hepatoma Huh-7 and Huh7.5.1–8 cell lines. *bioRxiv*, 2020.02.17.953281. doi:10.3389/fgene.2020.546106.
29. Wakita, T. *et al.* (2005) Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.*, 11, 791–796.
30. Kong, L. *et al.* (2020) Surfeit 4 contributes to the replication of hepatitis C virus using double-membrane vesicles. *J. Virol.*, 94, e00858-19.
31. Kong, L. *et al.* (2016) Prolactin regulatory element binding protein is involved in hepatitis C virus replication by interaction with NS4B. *J. Virol.*, 90, 3093–3111.
32. Kanehisa, M. *et al.* (2019) New approach for understanding genome variations in KEGG. *Nucleic Acids Res.*, 47(D1), D590–D595.
33. Ferrari, F. *et al.* (2011) PREDA: an R-package to identify regional variations in genomic data. *Bioinformatics*, 27, 2446–2447.
34. Jeong, D. *et al.* (2014) Cyr61 expression is associated with prognosis in patients with colorectal cancer. *BMC Cancer*, 14, 164.
35. Shi, J. *et al.* (2019) CYR61, a potential biomarker of tumor inflammatory response in epithelial ovarian cancer microenvironment of tumor progress. *BMC Cancer*, 19, 1140.
36. Huang, Y.T. *et al.* (2017) The matricellular protein CYR61 promotes breast cancer lung metastasis by facilitating tumor cell extravasation and suppressing anoikis. *Oncotarget*, 8, 9200–9215.
37. Langfelder, P. *et al.* (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9, 559.
38. Ikeda, K. *et al.* (1999) Effect of interferon therapy on hepatocellular carcinogenesis in patients with chronic hepatitis type C: a long-term observation study of 1,643 patients using statistical bias correction with proportional hazard analysis. *Hepatology*, 29, 1124–1130.
39. Perez, S. *et al.* (2019) Hepatitis C virus leaves an epigenetic signature post cure of infection by direct-acting antivirals. *PLoS Genet.*, 15, e1008181.
40. Perbal, B. (2004) CCN proteins: multifunctional signalling regulators. *Lancet*, 363, 62–64.
41. Li, Z.Q. *et al.* (2018) CCN1/Cyr61 enhances the function of hepatic stellate cells in promoting the progression of hepatocellular carcinoma. *Int. J. Mol. Med.*, 41, 1518–1528.
42. Li, Z.Q. *et al.* (2012) Cyr61/CCN1 is regulated by Wnt/β-catenin signaling and plays an important role in the progression of hepatocellular carcinoma. *PLoS One*, 7, e35754.
43. Makino, Y. *et al.* (2018) CTGF mediates tumor-stroma interactions between hepatoma cells and hepatic stellate cells to accelerate HCC progression. *Cancer Res.*, 78, 4902–4914.
44. Maity, G. *et al.* (2019) CYR61/CCN1 regulates dCK and CTGF and causes gemcitabine-resistant phenotype in pancreatic ductal adenocarcinoma. *Mol. Cancer Ther.*, 18, 788–800.
45. Young, N. *et al.* (2009) Sphingosine-1-phosphate regulates glioblastoma cell invasiveness through the urokinase plasminogen activator system and CCN1/Cyr61. *Mol. Cancer Res.*, 7, 23–32.
46. Aizarani, N. *et al.* (2019) A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature*, 572, 199–204.