

To determine the nucleotide sequence of the NS5A 2209–2248 region, we amplified nucleotides (nt) 7296–7320 of HCV complementary DNA by using the outer pair of primers [5' outer primer, 5'-TGG ATG GAG TGC GGT TGC ACA GGT A-3' (nt 6703–6727 of HC-J4); 3' outer primer, 5'-TCT TTC TCC GTG GAG GTG GTA TTG C-3' (nt 7296–7320)]. We transferred 1  $\mu$ l of the first PCR product to the second PCR reaction along with the nested 5' and 3' primers [5' inner primer, 5'-TGT AAA ACG ACG GCC AGT CAG GTA CGC TCC GGC GTG CA-3' (nt 6722–6741), with the M13 forward primer sequence underlined; and 3' inner primer, 5'-CAG GAA ACA GCT ATG ACC GGC GCC TTG GTA GGT GGC AA-3' (nt 7275–7294), with the M13 reverse primer sequence underlined]. An M13 forward primer and an M13 reverse primer were attached to the 5' terminal of the 5' and 3' inner primers, respectively, to facilitate direct sequencing with an automated DNA sequencer (model 373S; Applied Biosystems Japan).

Both strands of the PCR products were sequenced with the PRISM dye termination kit (Applied Biosystems Japan), according to the manufacturer's instructions. The sequencing primer was the M13 forward primer for the sense strand and the M13 reverse primer for the antisense strand. Deduced aa sequences of NS5A 2209–2248 were compared with the NS5A 2209–2248 sequences of HCV-J [15], which are prototypic sequences of HCV-1b. The results of the sequencing analysis were confirmed as consistent for each sample by repeating the experiment twice with different PCR products, to rule out the possibility of selection and amplification of minor NS5A quasi species variants in the low-titer specimens.

#### Nucleotide sequencing of the core gene

Substitutions of amino acids 70 and 91 in HCV-core region were determined according to core sequences obtained as described previously [16, 17]. The pattern of glutamine/histidine (mutant) at aa 70 and methionine (mutant) at aa 91 was evaluated as the double-mutant (dM) type, while the other patterns were non-double-mutant (non dM) type. Two patterns of mutants and competitive were labeled as non-wild. Wild at aa 70 and wild at aa 91 were evaluated as double-wild-type (dW), while the other patterns were considered non-double-wild-type (non dW).

#### Study design and treatment regimens

Patients were treated with combination therapy with PEG-IFN (Peg-Intron; Schering-Plough Nordic Biotech, Stockholm, Sweden) 1.2–1.5  $\mu$ g/kg subcutaneously and RBV (Rebetol; Schering-Plough Nordic Biotech) (body weight [b.w.]  $\leq$  60 kg, 600 mg po daily; b.w. 60–80 kg, 800 mg

po daily; b.w.  $>$  80 kg, 1000 mg po daily; in two divided doses). The duration of the combination therapy was set at a standard 48 weeks. Treatment reduction was permitted, to escape side effects, but extended treatment of 72 weeks is not included in this analysis. Achieved rates of PEG-IFN and RBV administration were calculated as the percentage of the actual total dose administered of a standard total dose of 48 weeks according to body weight before therapy. During treatment, patients were assessed as outpatients at weeks 2, 4, 6, and 8, and then every 4 weeks for the duration of treatment and at every 4 weeks after the end of therapy. Biochemical and hematological testing was done by a central laboratory. Serum HCV RNA was measured before treatment, during treatment at 4-weekly intervals, and after therapy at 4-weekly intervals for 24 weeks, by a quantitative PCR assay with a sensitivity of 100 copies/ml (National Genetics Institute, Los Angeles, CA, USA).

#### Outcomes

The primary end point was a sustained biochemical and virological response. Sustained virological response (SVR) was defined as serum HCV RNA undetectable at 24 weeks after the end of treatment. Secondary end points were end-of-treatment virological responses (HCV RNA undetectable in serum). In addition, tolerability (adverse events) and drug adherence were recorded and factors potentially associated with virological response were explored.

#### Statistical analysis

SPSS software package (SPSS 12J for Windows; SPSS, Chicago, IL, USA) was used for statistical analysis, which was carried out using the  $\chi^2$  or Fisher's exact probability test. Distributions of continuous variables were analyzed by the Mann-Whitney *U*-test. Independent factors possibly affecting response to combination therapy were examined by stepwise multiple logistic-regression analysis. All *P* values were two-tailed and those less than 0.05 were considered statistically significant.

## Results

#### Clinical characteristics and response to therapy

The clinical characteristics of the 239 patients are summarized in Table 1. On an intention-to-treat (ITT) analysis, serum HCV RNA levels were undetectable by the end of treatment in 172 of the 239 patients (72%) who were treated with PEG-IFN plus RBV, and among them, 98 of the 239 patients (41%) had an SVR (Table 2). The SVR rate decreased with drug discontinuation and dose

**Table 1** Baseline characteristics of participating patients infected with HCV genotype 1b

Total number	239
Age (years) <sup>a</sup>	57 (21–78)
Gender (male/female)	142/97
Body mass index (kg/m <sup>2</sup> ) <sup>a</sup>	23.3 (15.3–31.0)
Previous interferon therapy (no/yes)	167/72
Histology at biopsy	
Grade of inflammation	
A0/1/2/3	3/65/102/10
Stage of fibrosis	
F0/1/2/3/4	4/73/57/37/9
Hemoglobin (g/dl) <sup>b</sup>	14.3 ± 1.3
ALT (IU/L) <sup>b</sup>	86 ± 67
Platelet count (× 10 <sup>3</sup> /μl) <sup>b</sup>	160 ± 58
LDL cholesterol (mg/dl) <sup>b</sup>	74 ± 19
Serum HCV-RNA level (Log(IU/ml)) <sup>b, c</sup>	6.1 ± 0.6
Type of mutations in the core (dM/non dM)	30/166
Type of mutations in the core (dW/non dW)	65/131
Type of ISDR sequence (0/1/2/3/4 or more)	126/45/11/5/18

HCV hepatitis C virus, LDL low density lipoprotein, ALT alanine transaminase, ISDR interferon sensitivity determining region in NS5A 2209–2248, dM double mutant: dual substitutions at amino acids 70 and 91, non dM non-double mutant: wild type or substitution at either amino acid 70 or 91, dW double wild: wild type at amino acids 70 and 91, non dW non-double wild: dual or substitution at either amino acid 70 or 91

<sup>a</sup> Median (range) values are shown

<sup>b</sup> Data are mean ± SD

<sup>c</sup> Data are shown as Log(IU/ml)

reduction. The SVR rates of patients who received a total cumulative treatment dose of PEG-IFN of more than 80% were almost twice as high as the rates of patients who received less than 80% (56%, 26%, and 9% with [80%, 60%–80% and <60% of the PEG-IFN dose,  $P < 0.001$ ). The SVR rates did not decrease with RBV reduction, as long as the cumulative treatment dose of RBV was more than 60%, but when the RBV reduction fell below 60%, the SVR rates were significantly lower (56%, 38%, and 10% with [80%, 60%–80%, and <60% of the RBV dose,  $P < 0.001$ ).

#### Factors associated with sustained virological response

Seven parameters that influenced the SVR rate were identified by univariate analysis, including stage of fibrosis at liver biopsy, hemoglobin, platelet count, serum HCV RNA level, the type of ISDR sequence, and adherence to PEG-IFN plus RBV (Table 3). On the other hand, the SVR rate was not related to gender ( $P = 0.07$ ), age or BMI. The amino acid substitution pattern was not significant in the overall analysis, but female patients with dual substitutions

**Table 2** Sustained response rates to treatment according to drug adherence

Characteristic	Number/total number (%)
Overall	
End of treatment	172/239 (72)
End of follow up	98/239 (41)
PEG-interferon- $\alpha$ 2b adherence	
End of treatment	
[80%	131/154 (85)
60–80%	19/27 (70)
<60%	22/58 (38)
End of follow up	
[80%	86/154 (56)
60–80%	7/27 (26)
<60%	5/58 (9)
Ribavirin adherence	
End of treatment	
[80%	113/134 (84)
60–80%	37/46 (80)
<60%	22/59 (37)
End of follow up	
[80%	74/133 (56)
60–80%	18/47 (38)
<60%	6/59 (10)

PEG pegylated

at amino acids 70 and 91 had a low tendency to achieve SVR. As shown in Table 4, gender differences existed in the mutations in ISDR and core regions based on therapeutic responses. Because there were rather fewer female than male patients, the type of ISDR sequence did not significantly influence the SVR in females. We also analyzed types of mutations in the core, and the amino acid substitution pattern was not significant in the male patients, but female patients with dual substitutions at amino acids 70 and 91 had a low tendency to achieve an SVR, as mentioned above. We also compared results between treatment-naïve patients and those who had failed previous IFN therapy (Table 5). As there were some differences in stage of fibrosis, platelet count, grade of inflammation, and gender in univariate analysis, treatment was comparably effective in both groups.

Finally we performed multivariate analysis in subjects with good drug adherence (Table 6), which identified only one parameter that influenced the SVR rate independently by variable selection: the number of mutations in the ISDR sequence (two or more: odds ratio [OR] = 5.181,  $P < 0.05$ ). This regression model was always obtained regardless of the variable selection method used, including conditional parameter estimation, Wald statistic, and

**Table 3** Clinical and virological characteristics of 239 patients treated with PEG-IFN plus RBV therapy, based on therapeutic response

	SVR (n = 98)	Non-SVR (n = 141)	P value
Age (years) <sup>a</sup>	56 (27–69)	58 (23–72)	NS
Gender (male/female)	65/33	77/64	0.070
Previous interferon therapy (no/yes)	68/30	99/42	NS
Grade of inflammation (A0–1/2–3)	31/50	37/62	NS
Stage of fibrosis (F0–2/3–4)	68/13	67/33	0.009
Body mass index (kg/m <sup>2</sup> ) <sup>a</sup>	23.3 (15.5–28.1)	23.3 (15.3–31.0)	NS
Pretreatment Hemoglobin (g/dl) <sup>b</sup>	14.6 ± 1.1	14.0 ± 1.4	\ 0.001
Pretreatment ALT (IU/ml) <sup>b</sup>	87 ± 68	86 ± 67	NS
Pretreatment platelet count (× 10 <sup>3</sup> /μl) <sup>b</sup>	178 ± 63	148 ± 51	\ 0.001
Pretreatment LDL cholesterol (mg/dl) <sup>b</sup>	78 ± 21	72 ± 18	NS
Pretreatment serum HCV-RNA level (Log(IU/ml)) <sup>b, c</sup>	5.9 ± 0.7	6.2 ± 0.4	\ 0.001
No. of mutations in the ISDR (0–1/2 or more)	66/23	105/11	0.002
Type of mutations in the core (dM/non dM)	9/76	21/90	NS
Type of mutations in the core (dW/non dW)	31/54	34/77	NS
PEG-interferon adherence ([ 80/60–80/\ 60%])	85/7/6	68/20/53	\ 0.001
Ribavirin adherence ([ 80/60–80/\ 60%])	72/19/7	60/28/53	\ 0.001

IFN interferon, RBV ribavirin, SVR sustained virological response, NS not significant, ALT alanine transaminase, ISDR interferon sensitivity determining region in NS5A<sub>2209–2248</sub>, core substitution of amino acids 70 and 91, dM double mutant: dual substitutions at amino acids 70 and 91, non dM non-double mutant: wild type or substitution at either amino acid 70 or 91, dW double wild: wild type at amino acids 70 and 91, non dW non-double wild: dual or substitution at either amino acid 70 or 91

<sup>a</sup> Median (range) values are shown

<sup>b</sup> Data are mean ± SD

<sup>c</sup> Data are shown as Log(IU/ml)

**Table 4** Mutations in the ISDR and core regions analyzed separately for gender based on therapeutic response

	SVR (n = 98)	Non-SVR (n = 141)	P value
No. of mutations in the ISDR (0–1/2 or more)			
Male	36/21	56/8	0.002
Female	30/2	49/3	NS
Type of mutations in the core (dM/non dM)			
Male	8/46	11/48	NS
Female	1/30	10/42	0.026
Type of mutations in the core (dW/non dW)			
Male	18/36	16/43	NS
Female	13/18	18/34	NS

likelihood ratio statistic in combination with forward or backward variable selection methods.

Comparison of SVR rates according to the number of mutations in the ISDR sequence

We analyzed first the percentage of patients with more than two mutations in the ISDR among 762 patients who received IFN therapy between December 2000 and April

2008 at Tokyo Medical and Dental University Hospital and associated hospitals. The percentage of patients with more than two mutations in the ISDR was between about 20% and 30% for all ages (Fig. 1a).

Secondly, we analyzed responses to PEG-IFN plus RBV treatment and serum levels of HCV RNA in relation to the number of mutations in the ISDR. In Fig. 1b, patients with SVR are indicated by open circles and those with non-SVR, by closed circles. Although the rate of SVR tended to be higher in patients with increasing numbers of mutations in the ISDR, 5 patients with more than two mutations in the ISDR who experienced drug discontinuation and dose reduction resulted in non-SVR.

We confirmed changes over time in VR rates in patients treated with PEG-IFN plus RBV (Fig. 1c). Patients with more than two mutations in the ISDR are indicated in the figure by open circles and those with none or one mutation in the ISDR, by closed circles. The VR rates tended to be high early in the treatment in patients with more than two mutations in the ISDR.

Finally we compared the PEG-IFN plus RBV treatment efficacy in two groups, divided based on ISDR mutations. Patients with more than two mutations in the ISDR had a significantly higher tendency to achieve SVR in both ITT and per-protocol (PP) analyses ( $P \leq 0.01$ ) (Fig. 1d), and

**Table 5** Clinical and virological characteristics of 239 patients treated with PEG-IFN plus RBV therapy, based on previous interferon therapy

Previous interferon therapy	No (n = 167)	Yes (n = 72)	P value
Sustained response rates	68/167 (41)	30/72 (42)	NS
Age ( $\geq 65$ )	127/40	57/15	NS
Gender (male/female)	93/74	49/23	0.074
Grade of inflammation (A0–1/2–3)	55/72	13/40	0.018
Stage of fibrosis (F0–2/3–4)	103/24	32/21	0.003
Pretreatment hemoglobin ( $\geq 14.5$ )	93/74	41/31	NS
Pretreatment platelet count ( $\geq 160 \times 10^3$ )	84/83	50/22	0.006
Pretreatment Serum HCV RNA level <sup>a</sup> ( $\geq 6$ )	54/112	25/46	NS
No. of mutations in the ISDR (0–1/2 or more)	116/22	55/12	NS
PEG-interferon adherence ( $\geq 80$ – $\geq 80$ – $\geq 60$ %)	110/18/39	43/9/20	NS
Ribavirin adherence ( $\geq 80$ – $\geq 80$ – $\geq 60$ %)	97/30/40	35/17/20	NS

<sup>a</sup> Data are shown as Log(IU/ml)

**Table 6** Multivariate analysis for the clinical and virological factors related to sustained response to PEG-IFN plus RBV therapy in 104 patients who were not intolerant to PEG-IFN plus RBV therapy

Factor	Category	Odds ratio (95% CI)	P value
(a) Five-factor model			
Number of mutations in the ISDR	0 or 1	1	0.063
	2 or more	4.486 (0.922–21.74)	
Pretreatment Hemoglobin (g/dl)		1.250 (0.853–1.833)	NS
Pretreatment Serum HCV RNA level <sup>a</sup>		0.510 (0.224–1.159)	NS
Stage of fibrosis	F 0/1/2	1	NS
	F 3/4	0.460 (0.153–1.382)	
Pretreatment Platelet count ( $\times 10^3/\mu\text{l}$ )		1.022 (0.949–1.101)	
(b) Step-wise variable selection			
Number of mutations in the ISDR	0 or 1	1	0.034
	2 or more	5.181 (1.129–23.81)	

CI confidence interval, ALT alanine transaminase, ISDR interferon sensitivity determining region in NSSA 2209–2248

<sup>a</sup> Data are shown as Log(IU/ml)

the SVR rates of the patients with good drug adherence was 80%.

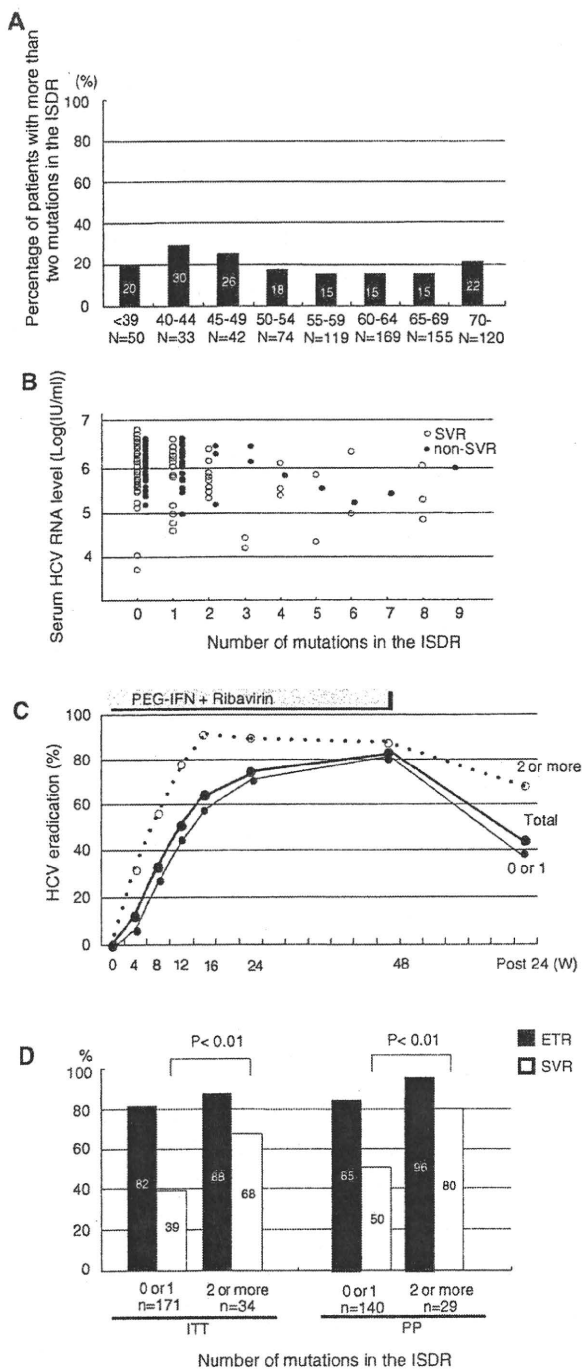
#### Side effects

Side effects leading to treatment discontinuation occurred in 53 patients (22%). Overall, 109 patients (46%) required reduction of the dose of one or both drugs during the treatment regimens (23% required PEG-IFN reduction and 35% required RBV reduction). The most common events leading to drug withdrawal were general fatigue and appetite loss ( $n = 15$ ), hematologic abnormalities ( $n = 6$ ), dermatological symptoms ( $n = 5$ ), retinopathy ( $n = 5$ ), neuro-psychiatric events ( $n = 4$ ), and interstitial pneumonia, including severe cough ( $n = 4$ ).

#### Discussion

Although the relationship between ISDR mutations and the clinical efficacy of IFN has been conflicting in Western countries [18–24], our results support previous studies reporting a close correlation between the number of mutations in the ISDR and IFN efficacy in patients with chronic HCV-1b infection [11–13]. Because most patients with 4 or more mutations in the ISDR (hereafter classified as the mutant type) experienced SVR with conventional IFN monotherapy, we reported previously that the number of amino acid substitutions in the ISDR was an independent predictor of the response to IFN therapy [12]. In the present study, we demonstrate that ISDR mutations are the most effective predictors of treatment outcome of 48-week





**Fig. 1** a The percentages of patients with more than two mutations in the interferon sensitivity determining region in NS5A 2209–2248 (ISDR), according to age (horizontal axis) among 762 patients who received interferon (IFN) therapy between December 2000 and April 2008 at Tokyo Medical and Dental University Hospital and associated hospitals. b Responses to pegylated (PEG)-IFN plus ribavirin (RBV) treatment and serum levels of hepatitis C virus (HCV) RNA in relation to the number of mutations in the ISDR. Patients with sustained virological response (SVR) are indicated by open circles and those with non-SVR by closed circles. c Changes over time in VR rates in patients treated with PEG-IFN plus RBV. Patients with more than two mutations in the ISDR are indicated by open circles and those with no or one mutation in the ISDR by closed circles, W weeks. d PEG-IFN plus RBV treatment efficacy divided into two groups based on ISDR mutations. End-of-treatment response (ETR) and SVR are shown in both intention-to-treat (ITT) analysis (left) and per-protocol (PP) analysis (right)

regard to age, there was no relation to SVR in overall analysis with continuous variables, but younger patients, aged less than 65 years, had a higher rate of response than those aged more than 65 years ( $P < 0.05$ , data not shown). Actually there are some reports suggesting the relationship of age and SVR [25, 26]. Finally, in regard to previous IFN therapy, as shown in Table 5, treatment was comparably effective in both groups; previous IFN therapy did not affect the SVR rate. The reasons for equivalent response rates in subjects with prior IFN history, which was not expected, are unclear. In our study, the group with prior IFN history had more advanced liver fibrosis and a low platelet count, and stage of fibrosis was one of the factors extracted by univariate analysis as a useful pretreatment marker predicting SVR. We also analyzed the other three parameters extracted by univariate analysis. Although there was no difference in pretreatment hemoglobin, or number of ISDR mutations, the group with prior IFN history tended to have a low serum HCV-RNA level. Further, the group with prior IFN history had a high proportion of male patients. Although the SVR rate was not related to gender, male subjects had a higher tendency to achieve SVR than female subjects.

In our present study, the SVR rate was not related to core mutations. As described in previous reports [17, 27, 28], amino acid substitutions in the core region are regarded as predictors of response to PEG-IFN plus RBV therapy in Japanese patients infected with HCV genotype 1b. In the present study, the SVR rate was not related to the pattern of amino acid substitution in the overall analysis. The reasons for these discrepant results are unclear, but females with dual substitutions at amino acids 70 and 91 had a lower tendency to achieve SVR. Further studies are necessary to clarify the mechanism of action for amino acid substitutions in the core region of HCV.

Recent studies suggest that the mutations in the ISDR are associated with response to combination therapy with IFN and RBV [29–32]. Most recently, it has been reported

**PEG-IFN plus RBV therapy in patients with HCV genotype 1b infection.**

In the present study, the SVR rate was not related to gender, age, or previous IFN therapy by univariate analysis. First of all, in regard to gender ( $P = 0.07$ ), as male patients had a higher tendency to achieve SVR than female patients, further validation in larger-scale studies is required to clarify the significance of gender. Secondly, in

that amino acid substitutions in the core and mutations in the ISDR are predictive of virological response to the combination therapy in patients with HCV genotype 1b and a high viral load [28]. There are some reports suggesting that the mutations in the ISDR may not serve as a predictor for treatment outcome [33, 34], but as the numbers of subjects in these studies were around 30, a number which is not sufficient to evaluate the results, this factor may explain these discrepant results.

The mechanisms of IFN sensitivity in relation to the sequence of the HCV NS5A<sub>2209–2248</sub> region are not clear. However the “mutant-type” ISDR correlates with a low viral load, as reported previously [12, 35, 36]; most patients in the present study with two or more mutations in the ISDR had high levels of virus. Furthermore, stepwise multiple logistic regression analysis of the factors, including substitution of the ISDR and the viral load, revealed that both of them were independent predictive variables of SVR, and the odds ratio of the number of mutations in the ISDR was the highest in the pretreatment factors associated with SVR by multivariate analysis. The precise mechanism involved must be elucidated in further *in vitro* studies.

There have been several reports that suggest biological roles of the ISDR in the response to IFN and in HCV infection. Double-stranded RNA-dependent protein kinase (PKR) is a critical component of the cellular antiviral responses induced by IFN. Gale et al. [37, 38] have reported that mutations within the PKR-binding region of NS5A, including ISDR, can disrupt the NS5A–PKR interaction, possibly rendering HCV sensitive to the antiviral effects of IFN. Toll-like receptor (TLR) has also been reported to play various roles in many viral infections, and it has been reported that NS5A bound MyD88, a major adaptor molecule of TLR-mediated signaling, and inhibited the TLR–MyD88 signaling pathway by a direct interaction with the death domain of MyD88 through the ISDR [39]. Furthermore, it has been reported that the lipid droplet is an important organelle for HCV production, and NS5A is a key protein that recruits replication complexes to lipid droplets for the production of infectious viral particles [40]. While the mechanism of action of the ISDR in the response to IFN or viral replication remains to be proven, these findings suggest new aspects of HCV infections.

In our previous report [12], patients with 4 or more mutations in the ISDR experienced SVR with conventional IFN monotherapy, but in more effective therapy with PEG-IFN plus RBV combination therapy, the number of mutations as a predictor of SVR decreased from 4 to 2. Watanabe et al. [41] have also reported that the number and position of mutations in the ISDR correlated with IFN efficacy in HCV-1b infection. Moreover, it has been reported that patients with viruses mutated at

positions 2209, 2216, or 2227 more frequently experienced SVR than did those without these mutations. Another group has also reported regarding statistical analysis, using a database of 675 individual ISDR sequences in HCV-NS5A and the IFN response [42]. They have shown that IFN-sensitive viruses contain a larger and more diverse collection of substitutions than IFN-resistant viruses. While it remains unknown how the numbers of mutations are involved in the biological role of ISDR, or which sites of mutation and changes of amino acid are also important for the response to IFN-based treatment, it is thought that the functional importance of numbers or sites of mutations can be explained in terms of interaction between NS5A and some target molecules such as PKR, MyD88, and lipid droplets.

*In vitro* studies have shown that the introduction of NS5A mutations enables an HCV replicon to replicate efficiently [10, 43, 44]. In our previous report, site-specific mutation of the ISDR also modulated HCV replication [45]. The ISDR was identified originally as the site that determines the sensitivity of HCV to IFN [12]. This indicates that the ISDR mutations are not lethal *in vivo*. Furthermore, mutations in the ISDR are closely associated clinically with decreased serum HCV RNA levels [42], whereas ISDR mutations in the HCV replicon enhance replication. While the explanation for this paradox has not become clear, a big difference between the environment of cultured cells and that in the human liver is thought contribute to this phenomenon.

We found that the percentage of patients with more than two mutations in the ISDR was between 20% and 30% for all ages; thus, around one-fifth of patients are thought likely to experience SVR. Indeed, the SVR rate among patients with two or more mutations in the ISDR sequence was 68% (ITT) and 80% (PP) compared to 39% (ITT) and 50% (PP) among those patients with no or one mutation in the present study. Furthermore, predictive factors such as serum HCV RNA level, stage of fibrosis, and hemoglobin also aid in the assessments of treatment, and we can use these parameters to develop a treatment strategy.

Several prospective randomized trials have shown that 72-week extended therapy improves SVR by 7.5%–12% in late viral responders [46, 47]. One cohort study showed that 72-week treatment for late viral responders achieved an even higher SVR, of 67.1%, which was 21% higher than the SVR achieved with 48-week treatment [48]. These reports demonstrate that tailoring of treatment duration by on-treatment viral response can further improve the outcomes of antiviral therapy. In our 48-week based treatment, 90% of patients with more than 2 ISDR mutations cleared the virus within 12 weeks of treatment (early viral response; EVR) and consequently achieved 30% higher SVR than those with 1 or no ISDR mutation. These results

suggest that ISDR mutations will remain a significant predictor of good response to IFN therapies, including 72-week extension.

In conclusion, ISDR mutations are the most effective predictors of treatment outcomes in multivariate analysis. The number of mutations in the ISDR sequence of HCV-1b ( $\geq 2$ ) is the most effective parameter which will facilitate further the selection of patients with a high likelihood of response to PEG-IFN plus RBV treatment.

**Acknowledgments** This study was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology-Japan; the Japan Society for the Promotion of Science; the Ministry of Health, Labour and Welfare-Japan; the Japan Health Sciences Foundation; the Miyakawa Memorial Research Foundation; and the National Institute of Biomedical Innovation. The following hospitals participated in the Ochanomizu-Liver Conference Study Group: Oume City General Hospital, Kashiwa City Hospital, Kudanzaka Hospital, Showa General Hospital, Shuwa General Hospital, Soka Municipal Hospital, Tama-Nambu Chiiki Hospital, Tsuchiura Kyodo General Hospital, Tokyo Kyosai Hospital, Tokyo Metropolitan Ohtsuka Hospital, Tokyo Metropolitan Fuchu Hospital, Tokyo Metropolitan Bokutoh Hospital, Toride Kyodo General Hospital, Nakano General Hospital, Hokushin General Hospital, Mishima Social Insurance Hospital, Musashino Red Cross Hospital, Yokosuka Kyosai Hospital, Yokohama City Minato Red Cross Hospital.

## References

1. Major ME, Feinstone SM. The molecular virology of hepatitis C. *Hepatology*. 1997;25:1527–38.
2. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*. 1989;244:359–62.
3. Reichard O, Andersson J, Schvarcz R, Weiland O. Ribavirin treatment for chronic hepatitis C. *Lancet*. 1991;337:1058–61.
4. Dibisceglie AM, Shindo M, Fong TL, Fried MW, Swain MG, Bergasa NV, et al. A pilot-study of ribavirin therapy for chronic hepatitis-C. *Hepatology*. 1992;16:649–54.
5. Dusheiko G, Main J, Thomas H, Reichard O, Lee C, Dhillon A, et al. Ribavirin treatment for patients with chronic hepatitis C: results of a placebo-controlled study. *J Hepatol*. 1996;25:591–8.
6. Reichard O, Norkrans G, Fryden A, Braconier JH, Sonnerborg A, Weiland O, et al. Randomised, double-blind, placebo-controlled trial of interferon alpha-2b with and without ribavirin for chronic hepatitis C. *Lancet*. 1998;351:83–7.
7. Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, et al. Randomised trial of interferon alpha 2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha 2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. *Lancet*. 1998;352:1426–32.
8. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, et al. Interferon alpha-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. *N Engl J Med*. 1998;339:1485–92.
9. Glue P, Fang JWS, Rouzier-Panis R, Raffanel C, Sabo R, Gupta SK, et al. Pegylated interferon-alpha 2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. *Clin Pharmacol Ther*. 2000;68:556–67.
10. Fried MW, Shiffman ML, Reddy KR, Smith C, Marionos G, Goncalves FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med*. 2002;347:975–82.
11. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest*. 1995;96:224–30.
12. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med*. 1996;334:77–81.
13. Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. *Hepatology*. 1997;25:750–3.
14. Enomoto N, Kurosaki M, Tanaka Y, Marumo F, Sato C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. *J Gen Virol*. 1994;75:1361–9.
15. Kato N, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, et al. Molecular-cloning of the human hepatitis-C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc Natl Acad Sci USA*. 1990;87:9524–8.
16. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. *Intervirology*. 2005;48:372–80.
17. Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. *Hepatology*. 2007;46:1357–64.
18. Zeuzem S, Lee JH, Roth WK. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon Alfa. *Hepatology*. 1997;25:740–4.
19. Khorsi H, Castelain S, Wyseur A, Izopet J, Canva V, Rombout A, et al. Mutations of hepatitis C virus 1b NS5A 2209–2248 amino acid sequence do not predict the response to recombinant interferon-alfa therapy in French patients. *J Hepatol*. 1997;27:72–7.
20. Squadrito G, Leone F, Sartori M, Nalpas B, Berthelot P, Raimondo G, et al. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alfa. *Gastroenterology*. 1997;113:567–72.
21. Hofgartner WT, Polyak SJ, Sullivan DG, Carithers RL, Gretch DR. Mutations in the NS5A gene of hepatitis C virus in North American patients infected with HCV genotype 1a or 1b. *J Med Virol*. 1997;53:118–26.
22. Squadrito G, Orlando ME, Cacciola I, Rumi MG, Artini M, Picciotto A, et al. Long-term response to interferon alpha is unrelated to “interferon sensitivity determining region” variability in patients with chronic hepatitis C virus-1b infection. *J Hepatol*. 1999;30:1023–7.
23. Chung RT, Monto A, Dienstag JL, Kaplan LM. Mutations in the NS5A region do not predict interferon-responsiveness in American patients infected with genotype 1b hepatitis C virus. *J Med Virol*. 1999;58:353–8.
24. Sarrazin C, Berg T, Lee JH, Teuber G, Dietrich CF, Roth WK, et al. Improved correlation between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. *J Hepatol*. 1999;30:1004–13.
25. Honda T, Katano Y, Urano F, Murayama M, Hayashi K, Ishigami M, et al. Efficacy of ribavirin plus interferon-alpha in patients aged  $\geq 60$  years with chronic hepatitis C. *J Gastroenterol Hepatol*. 2007;22:989–95.

26. Hung CH, Chen CH, Lee CM, Wu CM, Hu TH, Wang JH, et al. Association of amino acid variations in the NS5A and E2-PePHD region of hepatitis C virus 1b with hepatocellular carcinoma. *J Viral Hepatitis*. 2008;15:58–65.
27. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Predictive factors of virological non-response to interferon-ribavirin combination therapy for patients infected with hepatitis C virus of genotype 1b and high viral load. *J Med Virol*. 2006;78:83–90.
28. Mori N, Imamura M, Kawakami Y, Saneto H, Kawaoka T, Takaki S, et al. Randomized trial of high-dose interferon-alpha-2b combined with ribavirin in patients with chronic hepatitis C: correlation between amino acid substitutions in the core/NS5A region and virological response to interferon therapy. *J Med Virol*. 2009;81:640–9.
29. Yen YH, Hun CH, Hu TH, Chen CH, Wu CM, Vvang JH, et al. Mutations in the interferon sensitivity-determining region (non-structural 5A amino acid 2209–2248) in patients with hepatitis C-1b infection and correlating response to combined therapy of pegylated interferon and ribavirin. *Aliment Pharmacol Ther*. 2008;27:72–9.
30. Hung CH, Lee CM, Lu SN, Lee JF, Wang JH, Tung HD, et al. Mutations in the NS5A and E2-PePHD region of hepatitis C virus type 1b and correlation with the response to combination therapy with interferon and ribavirin. *J Viral Hepatitis*. 2003;10:87–94.
31. Murayama M, Katano Y, Nakano I, Ishigami M, Hayashi K, Honda T, et al. A mutation in the interferon sensitivity-determining region is associated with responsiveness to interferon-ribavirin combination therapy in chronic hepatitis patients infected with a Japan-specific subtype of hepatitis C virus genotype 1b. *J Med Virol*. 2007;79:35–40.
32. Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. *Hepatology*. 2008;48:1753–60.
33. Murphy MD, Rosen HR, Marousek GI, Chou SW. Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. *Dig Dis Sci*. 2002;47:1195–205.
34. Yang SS, Lai MY, Chen DS, Chen GH, Kao JH. Mutations in the NS5A and E2-PePHD regions of hepatitis C virus genotype 1b and response to combination therapy of interferon plus ribavirin. *Liver Int*. 2003;23:426–33.
35. Chayama K, Tsubota A, Kobayashi M, Okamoto K, Hashimoto M, Miyano Y, et al. Pretreatment virus load and multiple amino acid substitutions in the interferon sensitivity-determining region predict the outcome of interferon treatment in patients with chronic genotype 1b hepatitis C virus infection. *Hepatology*. 1997;25:745–9.
36. Watanabe H, Nagayama K, Enomoto N, Itakura J, Tanabe Y, Hamano K, et al. Sequence elements correlating with circulating viral load in genotype 1b hepatitis C virus infection. *Virology*. 2003;311:376–83.
37. Gale MJ, Korth MJ, Tang NM, Tan SL, Hopkins DA, Dever TE, et al. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. *Virology*. 1997;230:217–27.
38. Gale M, Blakely CM, Kwieciszewski B, Tan SL, Dossett M, Tang NM, et al. Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol*. 1998;18:5208–18.
39. Abe T, Kaname Y, Hamamoto I, Tsuda Y, Wen XY, Taguwa S, et al. Hepatitis C virus nonstructural protein 5A modulates the toll-like receptor-MyD88-dependent signaling pathway in macrophage cell lines. *J Virol*. 2007;81:8953–66.
40. Miyanari Y, Atsuzawa K, Usuda N, Watashi K, Hishiki T, Zayam M, et al. The lipid droplet is an important organelle for hepatitis C virus production. *Nat Cell Biol*. 2007;9:1089–97.
41. Watanabe H, Enomoto N, Nagayama K, Izumi N, Marumo F, Sato C, et al. Number and position of mutations in the interferon (IFN) sensitivity-determining region of the gene for nonstructural protein 5A correlate with IFN efficacy in hepatitis C virus genotype 1b infection. *J Infect Dis*. 2001;183:1195–203.
42. Witherell GW, Beineke P. Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. *J Med Virol*. 2001;63:8–16.
43. Blight KJ, Kolykhalov AA, Rice CM. Efficient initiation of HCV RNA replication in cell culture. *Science*. 2000;290:1972–4.
44. Maekawa S, Enomoto N, Sakamoto N, Kurosaki M, Ueda E, Kohashi T, et al. Introduction of NS5A mutations enables subgenomic HCV replicon derived from chimpanzee-infectious HC-J4 isolate to replicate efficiently in Huh-7 cells. *J Viral Hepatitis*. 2004;11:394–403.
45. Kohashi T, Maekawa S, Sakamoto N, Kurosaki M, Watanabe H, Tanabe Y, et al. Site-specific mutation of the interferon sensitivity-determining region (ISDR) modulates hepatitis C virus replication. *J Viral Hepatitis*. 2006;13:582–90.
46. Berg T, von Wagner M, Nasser S, Sarrazin C, Heintges T, Gerlach T, et al. Extended treatment duration for hepatitis C virus type 1: comparing 48 versus 72 weeks of peginterferon-alfa-2a plus ribavirin. *Gastroenterology*. 2006;130:1086–97.
47. Ferenci P, Laferl H, Scherzer TM, Maieron A, Hofer H, Stauber R, et al. Peginterferon alfa-2a/ribavirin for 48 or 72 weeks in hepatitis C types 1 and 4 patients with slow virologic response. *Gastroenterology* 2009 [Epub ahead of print].
48. Watanabe S, Enomoto N, Koike K, Izumi N, Takikawa H, Hashimoto E, et al. Prolonged treatment with pegylated interferon a 2b plus ribavirin improves sustained virological response in chronic hepatitis C genotype 1 patients with late response in a clinical real-life setting in Japan. *Hepatol Res* 2009 [Epub ahead of print].

## Comparison of HCV-associated gene expression and cell signaling pathways in cells with or without HCV replicon and in replicon-cured cells

Yuki Nishimura-Sakurai · Naoya Sakamoto · Kaoru Mogushi · Satoshi Nagaie · Mina Nakagawa · Yasuhiro Itsui · Megumi Tasaka-Fujita · Yuko Onuki-Karakama · Goki Suda · Kako Mishima · Machi Yamamoto · Mayumi Ueyama · Yusuke Funaoka · Takako Watanabe · Seishin Azuma · Yuko Sekine-Osajima · Sei Kakinuma · Kiichiro Tsuchiya · Nobuyuki Enomoto · Hiroshi Tanaka · Mamoru Watanabe

Received: 2 September 2009 / Accepted: 2 November 2009 / Published online: 12 December 2009  
© Springer 2009

### Abstract

**Background** Hepatitis C virus (HCV) replication is affected by several host factors. Here, we screened host genes and molecular pathways that are involved in HCV replication by comprehensive analyses using two genotypes of HCV replicon-expressing cells, their *cured* cells and naïve Huh7 cells.

Y. Nishimura-Sakurai and N. Sakamoto contributed equally to this work.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00535-009-0162-3) contains supplementary material, which is available to authorized users.

Y. Nishimura-Sakurai · N. Sakamoto (✉) · M. Nakagawa · Y. Itsui · M. Tasaka-Fujita · Y. Onuki-Karakama · G. Suda · K. Mishima · M. Yamamoto · M. Ueyama · Y. Funaoka · T. Watanabe · S. Azuma · Y. Sekine-Osajima · S. Kakinuma · K. Tsuchiya · M. Watanabe  
Department of Gastroenterology and Hepatology,  
Tokyo Medical and Dental University, 1-5-45 Yushima,  
Bunkyo-ku, Tokyo 113-8519, Japan  
e-mail: nsakamoto.gast@tmd.ac.jp

N. Sakamoto · M. Nakagawa · S. Kakinuma  
Department for Hepatitis Control,  
Tokyo Medical and Dental University, Tokyo, Japan

K. Mogushi · S. Nagaie · H. Tanaka  
Information Center for Medical Science,  
Tokyo Medical and Dental University, Tokyo, Japan

Y. Itsui  
Department of Internal Medicine,  
Soka Municipal Hospital, Saitama, Japan

N. Enomoto  
First Department of Internal Medicine,  
University of Yamanashi, Yamanashi, Japan

**Methods** Huh7 cell lines that stably expressed HCV genotype 1b or 2a replicon were used. The *cured* cells were established by treating HCV replicon cells with interferon-alpha. Expression of 54,675 cellular genes was analyzed by GeneChip DNA microarray. The data were analyzed by using the KEGG Pathway database.

**Results** Hierarchical clustering analysis showed that the gene-expression profiles of each cell group constituted clear clusters of naïve, HCV replicon-expressed, and cured cell lines. The pathway process analysis between the replicon-expressing and the *cured* cell lines identified significantly altered pathways, including MAPK, steroid biosynthesis and TGF-beta signaling pathways, suggesting that these pathways were affected directly by HCV replication. Comparison of *cured* and naïve Huh7 cells identified pathways, including steroid biosynthesis and sphingolipid metabolism, suggesting that these pathways were required for efficient HCV replication. Cytoplasmic lipid droplets were obviously increased in replicon-expressing and *cured* cells as compared to naïve cells. HCV replication was significantly suppressed by peroxisome proliferator-activated receptor (PPAR)-alpha agonists but augmented by PPAR-gamma agonists.

**Conclusion** Comprehensive gene expression and pathway analyses show that lipid biosynthesis pathways are crucial to support proficient virus replication. These metabolic pathways could constitute novel antiviral targets against HCV.

**Keywords** DNA microarray · KEGG database · HCV replicon · Lipid metabolism

### Abbreviations

HCV      Hepatitis C virus  
TLR      Toll-like receptor  
BMP      Bone morphogenetic protein



TGF	Transforming growth factor
FKBP	FK-binding protein
HSP	Heat shock proteins
FBS	Fetal bovine serum
YFP	Yellow fluorescence protein
FACS	Fluorescent activated cell sorting
RIN	RNA integrity number
SAM	Significance analysis of microarray
KEGG	Kyoto Encyclopedia of Genes and Genomes
EGID	NCBI Entrez Gene ID
RT-PCR	Reverse transcription-polymerase chain reaction
MTS	Dimethylthiazol carboxymethoxyphenyl sulfophenyl tetrazolium
PPAR	Peroxisome proliferator-activated receptor

## Introduction

Hepatitis C virus (HCV) infection is one of the most important causative agents of acute and chronic hepatitis, liver cirrhosis and hepatocellular malignancies [1]. Currently, the most efficient combination treatment of ribavirin plus peginterferon can eliminate the virus in almost half of the patients treated [2, 3]. Thus, it is our high priority goal to understand the HCV life cycle precisely, to identify cellular cofactors for HCV replication and to develop new class antiviral therapeutics.

Molecular analyses of the HCV life cycle, virus–host interactions, and mechanisms of liver cell damage by the virus are not understood completely, mainly because of the lack of cell culture systems. These problems have been partly overcome by the development of the HCV subgenomic replicon [4] and HCV cell culture systems [5, 6]. These systems have allowed us to study the complete HCV life cycle: virus–cell entry, translation, protein processing, RNA replication, virion assembly and virus release.

Several host proteins and drugs have been reported to have a direct effect on HCV replication *in vitro* [7]. These include factors that affect immune responses (interferons and their related genes [8, 9], RIG-I, TLRs [10]), cell proliferation (BMP7 [11], TGF- $\beta$  [12], nucleolin [13]), molecular chaperone function (cyclophilin [14], ER-stress proteins [15], FKBP [16], HSP27 [17], HSP90 [18]) and lipid metabolism (cholesterol, sphingolipid [19]). However, it is often difficult to determine whether these genes are changed by HCV replication or the changes are essential for HCV replication in the host cells.

In this study, we investigated the effects of host cellular gene expression using our HCV replicon system [20, 21]. We performed DNA microarray analyses using cells expressing the replicons, the corresponding *cured* cells, from which the replicon had been eliminated by prolonged treatment with interferon- $\alpha$ , and naïve Huh7 cells. Furthermore, we investigated the signaling pathways using DNA microarrays to study molecular pathways that are involved in the HCV life cycle and its pathogenesis.

## Materials and methods

### Cells and cell culture

Huh7 cells were maintained in Dulbecco's modified minimal essential medium (Sigma, St. Louis, MO) supplemented with 10% fetal calf serum at 37°C under 5% CO<sub>2</sub>. To maintain cell lines carrying the HCV replicon (Huh7/Rep cells), G418 (Nakalai Tesque, Kyoto, Japan) was added to the culture medium to a final concentration of 500  $\mu$ g/ml.

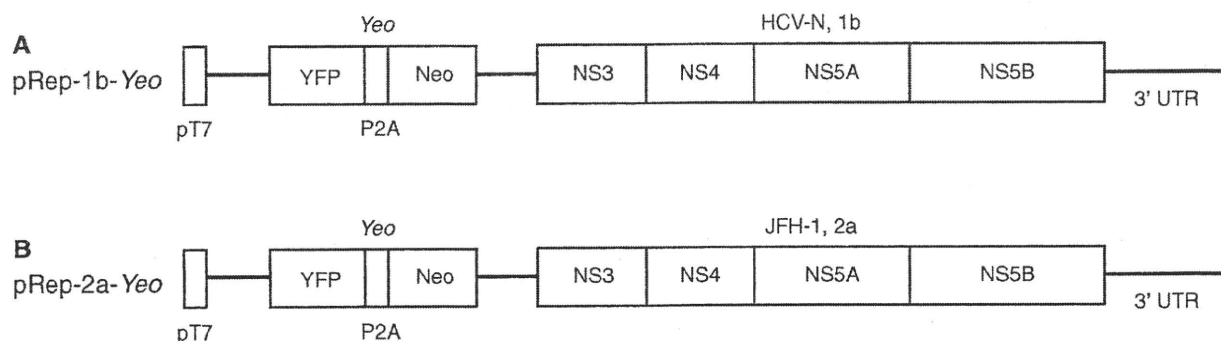
### HCV replicon and cell culture

The HCV-1b replicon plasmid, pHCV1bneo-delS, was provided by Dr. Christoph Seeger (Fox Chase Cancer Center, Philadelphia, PA) [22]. HCV-2a replicon plasmid, pSGR-JFH1, was provided by Dr. Takaji Wakita (National Institute of Infectious Diseases, Tokyo, Japan). The neomycin phosphotransferase (Neo) gene of pHCV1bneo-delS and pSGR-JFH1 was replaced by a chimeric gene coding for yellow fluorescent protein fused in-frame with the foot-and-mouth disease virus peptide 2A (P2A) autocleavage motif followed by neomycin phosphotransferase, which we designated *Yeo* (Fig. 1) [23]. An HCV *Feo*-replicon that expresses chimeric firefly luciferase and the neomycin resistance gene has been described [20, 21]. *In vitro* replicon RNA synthesis, RNA transfection and selection of G418-resistant cell lines were carried out as described previously [21, 24]. Briefly, replicon RNAs were transfected into Huh7 cells. By cell culture in the presence of G418, we established Huh7 cell lines that stably express the *Yeo*-replicons: Huh7/Rep-1b-*Yeo* and Huh7/Rep-2a-*Yeo*.

### Fluorescence microscopy and FACS analysis

The cells were plated onto eight-well chamber slides (Lab-Tek® Chamber Slide™ System, Nalgen Nunc International, Rochester, NY), and the YFP expression was detected by fluorescence microscopy (BZ-8000,





**Fig. 1** Structure of replicon plasmid constructs. A hepatitis C virus (HCV) replicon plasmid, pRep-1b-Yeo (**a**) and pRep-2a-Yeo (**b**), was reconstructed from pHCV1bneo-delS [22] and pSGR-JFH1 [47] by replacing the neomycin phosphotransferase (Neo) gene with a fusion

gene of yellow fluorescence protein (YFP) and Neo, which we designate “Yeo.” NS Nonstructural region, pT7 T7 promoter, 3'UTR 3'untranslated region, P2A foot-and-mouth disease peptide 2A (see also “Materials and methods”) [23]

KEYENCE, Osaka, Japan) and FACS Caliber using CellQuest software (BD Biosciences, Franklin Lakes, NJ).

#### Cell sorting

Cells were treated for 5 min with trypsin/EDTA at 37°C and then resuspended in 10% FBS/DMEM. A single cell suspension was prepared by passage through a 35- $\mu$ m nylon filter. The cell populations that support a high level of Yeo-replicon expression (Huh7/Rep-1b-Yeo<sup>high</sup> and Huh7/Rep-2a-Yeo<sup>high</sup>) were separated using a FACS Vantage SE cell-sorting system (BD Biosciences). The YFP-directed fluorescence of sorted cells was confirmed by fluorescence microscopy and FACS.

#### Establishment of the cured Huh7 cells

Cured Huh7 cells (cHuh7) were established by eliminating the HCV replicon from the Yeo-1b<sup>high</sup> and -2a<sup>high</sup> replicon expressing Huh7 cells by treatment with 100 U/ml of interferon-alpha for 14 days [6, 25]. Clearance of replicon RNA was confirmed by FACS analysis and by the loss of resistance to G418.

#### RNA preparation and microarray hybridization

Total cellular RNA was extracted from the 1b<sup>high</sup> and 2a<sup>high</sup> Yeo-replicon cells, cured-1b and -2a cells and naïve Huh7 cells using ISOGEN (Wako). Integrity of obtained RNA was assessed using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). All samples had an RNA Integrity Number (RIN) greater than 9.4 [26]. Complementary RNA was prepared from 1  $\mu$ g total RNA, using one-cycle target labeling and a control reagents kit (Affymetrix, Santa Clara, CA). Hybridization and signal detection of the Human Genome U133 Plus 2.0 array (Affymetrix) were performed in accordance with the

manufacturer's instructions. Assays were performed in duplicate.

#### Analysis of gene expression data

A total of ten microarray datasets was normalized using the robust multi-array average (RMA) method under R 2.8.1 statistical software (<http://www.R-project.org>). Estimated gene expression levels were  $\log_2$ -transformed, data from 62 control probe sets were removed, and we selected 18,613 probe sets that were categorized as “present” or “marginal” among all samples. We performed two sets of gene comparisons to examine effects of HCV replicons on host cellular gene expression: one was the high Yeo-replicon-expressing cells versus cured cells and the other was parental Huh7 cells versus cured cells. We selected differently expressed genes using the significance analysis of microarray (SAM) as described by Tusher et al. [27], and the fold changes and the *Q*-values were calculated for each probe sets. We used  $\Delta = 0.1$  as a cutoff parameter for SAM. A hierarchical clustering with selected genes was performed with R software. Euclidean distance was used to calculate the similarity matrix among genes or cell conditions, respectively. The complete linkage method was used for agglomeration.

#### Molecular pathway analysis and visualization of gene expression data

We used the KEGG Pathway database to investigate the molecular reactions and pathways that showed significant gene expression changes [28]. The KEGG Pathway is a database of biological systems, consisting of over 4,252 genes and 204 molecular pathway-wiring diagrams of interaction and reaction networks (<http://www.genome.jp/kegg/pathway.html>). Prior to the pathway analysis, we selected probe sets that were differentially expressed

between Huh7 and *cured* cells and between *cured* cells and replicon cells. For Huh7 versus *cured* cells analyses, we selected probe sets that showed 20% upregulation or downregulation (i.e., fold change of greater than 1.2) in both Huh7 versus *cured*-1b and Huh7 versus *cured*-2a cells. For replicon cell versus *cured* cell analyses, we selected probe sets that showed 20% upregulation or downregulation in both *cured*-1b versus replicon-1b cells or *cured*-2a versus replicon-2a cells. Association between the obtained gene list and each pathway was evaluated by Fisher's exact test. The significance level for KEGG analysis was set to a false discovery rate (FDR) of lower than 0.3 using the Benjamini and Hochberg method [29].

We next visualized functional associations between the differentially expressed genes and biological pathway processes. The KEGG Pathway provides a reference knowledge base for linking genomes to biological systems and also to environments by the processes of Pathway mapping and BRITe mapping. NCBI Entrez Gene IDs (EGIDs) for each gene in the pathways were extracted from the database. The relationship between probe sets on the microarray and EGIDs was obtained from a gene annotation file provided by Affymetrix. Thereafter, gene expression changes were mapped on the pathway by combining the results of fold-change analyses with the data sets above.

#### Real-time PCR analysis of mRNA expression

To confirm the results of the microarray analysis, we examined the expression levels of several mRNA by real-time RT-PCR (7500 Real Time PCR Systems, Applied Biosystems, Foster City, CA). Single-stranded cDNA was synthesized from total RNA using SuperScript II reverse transcriptase (Invitrogen) and random hexamers (Takara Bio Inc., Shiga, Japan) as primers. Expression of mRNA was quantified using QuantiTect SYBR Green PCR master Mix (QIAGEN, Valencia, CA). The primers used were as follows: HMGCR, SQLE, CYP51A1, TM7SF2, NSDHL, EBP and beta-actin. The nucleotide sequences of primers and corresponding product sizes are as indicated (see Supplementary Table 1).

#### Oil red O staining

Huh7 cells, replicon cells and *cured* cells were cultured on 18-mm-round micro cover glasses (Matsunami, Tokyo, Japan). These cells were fixed with 4% paraformaldehyde for 5 min at room temperature. After washing with PBS, the cells were permeabilized with 0.05% Triton X-100 in PBS for 5 min at room temperature. Staining of intracellular neutral lipids was performed with Oil red O, and nuclei were stained with Mayer's hematoxylin using Oil

red O stain kit procedure (Diagnostic Biosystems Inc., Pleasanton, CA).

#### Immunofluorescence analysis

Huh7 cells, replicon cells and *cured* cells were cultured on 18-mm-round micro cover glasses. For immunostaining, the cells were fixed in 4% paraformaldehyde for 5 min at room temperature. For detection of HCV-NS5A, cells were incubated with the primary antibody (Biosdesign International, Saco, ME) for 1 h at 37°C. The fluorescent secondary antibodies were Alexa Fluor 594 goat anti-mouse IgG antibody (Invitrogen, Carlsbad, CA). Nuclei were labeled with 4',6-diamidino-2-phenylindole (DAPI). Lipid droplets were visualized with BODIPY 493/503 (Invitrogen). Analysis was performed on a Delta-Vision microscope system (Applied Precision, Seattle, WA).

#### Luciferase-based expression analysis of HCV replicon and analysis of cell viability

Huh7/Rep-Feo cells [20, 21] were cultured with various concentrations of peroxisome proliferator-activated receptor (PPAR)-alpha and -gamma agonists. After 48 h of culture, levels of HCV replication were quantified by internal luciferase assay using a Bright-Glo Luciferase Assay System (Promega). Assays were performed in triplicate, and the results were expressed as mean  $\pm$  SD as percentages of the controls. To evaluate cell viability, dimethylthiazol carboxymethoxyphenyl sulfophenyl tetrazolium (MTS) assay was performed using a Cell Titer 96 Aqueous One Solution Cell Proliferation Assay (Promega) according to manufacturer's directions.

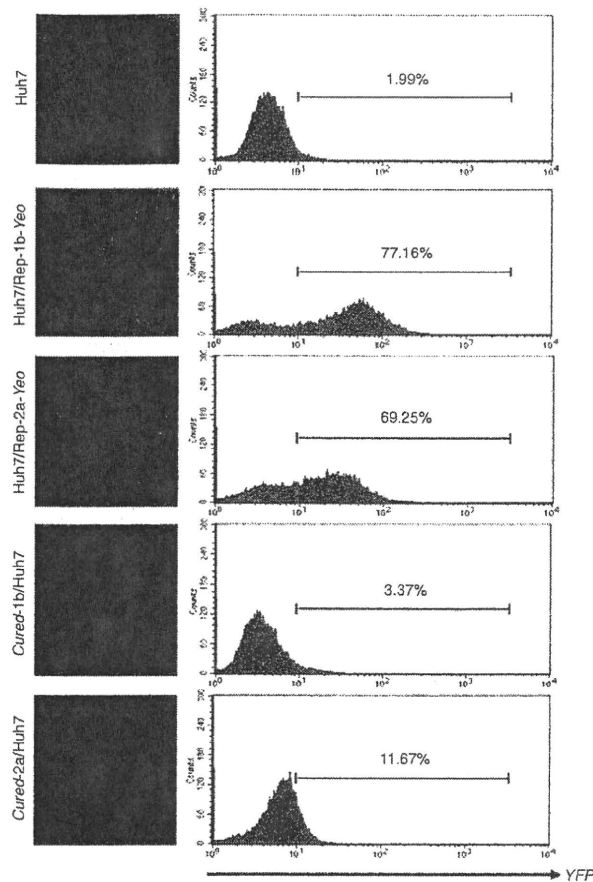
#### Statistical analyses

Statistical analyses were performed using the Student's *t*-test, and *P*-values of less than 0.05 were considered as statistically significant.

## Results

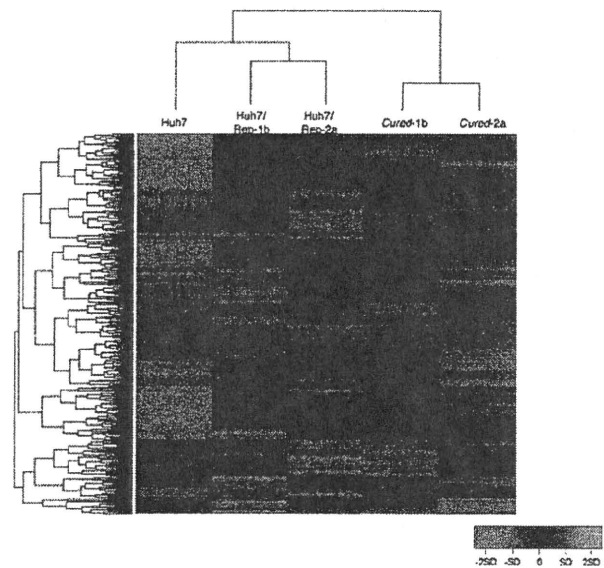
#### Fluorescence detection of Yeo replicon

Genotypes 1b and 2a *Yeo*-replicon RNAs were stably transfected into Huh7 cells (Huh7/Rep-1b-*Yeo* and Huh7/Rep-2a-*Yeo*, respectively, Fig. 1). In these transfected cells, expression of the HCV replicon was visualized by HCV-IRES-driven, YFP-mediated fluorescence (Fig. 2, left panels). The expression levels of individual cells could be measured by fluorescence intensity and cytogram analysis using flow cytometry (Fig. 2, right panels).



**Fig. 2** Visualization of YFP replicon expression. Detection of YFP expression by fluorescence microscopy analysis and of intracellular YFP expression by FACS analysis

Our initial trial was to compare gene expression profiles between replicon-expressing cells and parental Huh7 cells. However, these comparisons identified gene expressional changes induced by HCV infection and by adaptation of the host cell to support efficient HCV replication, because transfection of replicon RNA and G418-treatment of cells resulted in selection of a cell population that can support a high level of HCV subgenomic replication. Therefore, we used *cured* cell lines, which were established from Huh7/Rep-1b-Yeo and -2a-Yeo by interferon-alpha treatment. These *cured* cell lines are highly permissive for HCV replication on re-introduction of virus or replicon RNA (Fig. 2). With these backgrounds, we performed two sets of gene comparison using microarray analyses: comparison of replicon-expressing cell lines (Huh7/Rep-1b-Yeo and Huh7/Rep-2a-Yeo) and *cured* cells (Cured-1b/Huh7 and Cured-2a/Huh7) was intended to identify genes that are affected by HCV replication, and comparison of parental Huh7 cells and *cured* cell lines was intended to identify



**Fig. 3** Hierarchical clustering of gene expression profiles obtained from the 1b and 2a Yeo-replicon expressing cells, *cured* cells and Huh7. The 1,870 probe sets that changed expression more than 1.2-fold in either Huh7 versus *cured* or *cured* versus replicon were selected. Dendrograms show the classification determined by hierarchical clustering analysis. *Red* and *green* colors indicate relative overexpression and underexpression, respectively

genes that are essential for a high level of HCV replication in cultured cells.

Hierarchical clustering gene expression profiles in naïve, replicon-expressing and *cured* cells

Datasets from the microarrays were normalized using the robust multi-array average (RMA) method, and differentially expressed genes were extracted in replicon cells, *cured* cells and naïve Huh7 cells. Gene expression profiles were well correlated each other between duplicate microarray data from the same cell line with the Pearson's correlation coefficients ( $R^2$ ) of greater than 0.975 (see Supplementary Fig. 1). In this analysis, 1,870 probe sets showed differences in expression levels of more than 1.2-fold under  $\Delta < 0.1$  in either Huh7 versus *cured* (1,516 probe sets) or *cured* versus replicon (372 probe sets). A hierarchical clustering analysis showed that the gene-expression profiles of each cell group constituted clear clusters, Huh7/Rep-2a and Huh7/Rep-1b, Cured-2a and Cured-1b, and Huh7 cells (Fig. 3). Among genes whose expression differed significantly between replicon-expressing cells and *cured* cells, 15 showed changes of more than two-fold (Table 1). These included cell cycle- or cell growth-related genes (nuclear protein 1, growth differentiation factor 15, urothelial cancer associated 1, inhibin and tubulin), oncogene (Ras-related GTP binding D) and interferon-related gene (IFIM3). On the other hand, 37

**Table 1** Microarray analysis: genes for which the expression changed more than two-fold in Rep-1b/Huh7 and Rep-2a/Huh7 cells compared to their *cured* cells

Probe set	Title	Rep1b/cured1b		Rep2a/cured2a	
		Fold change	Q-value	Fold change	Q-value
209230_s_at	Nuclear protein 1	5.41	0.00	13.78	7.44
221523_s_at	Ras-related GTP binding D	3.82	0.00	2.32	24.60
221524_s_at	Ras-related GTP binding D	3.17	0.00	2.40	17.08
205923_at	Reelin	3.14	0.00	2.70	12.42
200924_s_at	Hypothetical protein LOC442497/solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	2.84	0.00	2.63	0.69
221577_x_at	Growth differentiation factor 15	2.62	0.00	3.16	0.00
233030_at	Patatin-like phospholipase domain containing 3	2.07	0.00	2.31	3.16
201471_s_at	Sequestosome 1	2.52	0.81	2.10	2.45
227919_at	Urothelial cancer associated 1	5.09	0.90	2.90	54.18
207076_s_at	Argininosuccinate synthetase 1	2.99	0.90	2.22	61.28
205749_at	Cytochrome P450, family 1, subfamily A, polypeptide 1	2.69	0.90	2.63	4.55
217127_at	Cystathionase (cystathionine gamma-lyase)	2.32	3.05	2.38	6.54
210587_at	Inhibin, beta E	2.51	4.25	3.89	2.45
214023_x_at	Tubulin, beta 2B	2.41	4.25	4.31	54.18
212203_x_at	Interferon induced transmembrane protein 3 (1-8U)	2.34	5.75	2.09	54.18

genes were up-regulated by more than two-fold between *cured* and naïve cells (Table 2), which included genes such as chemokine (CCL14), solute carrier family and metallothionein family.

#### Pathway process analyses and hierarchical clustering of genes in each functional category

Using the KEGG Pathway database, we analyzed pathway processes that were altered between replicon-expressing cells and *cured* cells as well as between *cured* cells and Huh7 cells (Supplementary Tables 2, 3). Comparison of the pathway processes between replicon-expressing and *cured* cells identified six pathways that showed differences of  $FDR < 0.3$ , including pathways related to MAPK ( $P = 4.0 \times 10^{-4}$ ,  $FDR 0.08$ ), biosynthesis of steroids ( $P = 4.21 \times 10^{-3}$ ,  $FDR 0.21$ ) and TGF-beta ( $P = 8.4 \times 10^{-3}$ ,  $FDR 0.29$ ) (KEGG Pathway maps for each significant pathways are shown in Supplementary Fig. 2A–F). Comparison of the pathway processes between *cured* and naïve Huh7 cells identified 11 significant pathways (KEGG Pathway maps for each significant pathway are shown in Fig. 5 and Supplementary Fig. 3A–J). These included pathways that were related to TGF-beta ( $P = 8.42 \times 10^{-3}$ ), cell cycle ( $P = 9.0 \times 10^{-3}$ ) and sphingolipid metabolism ( $P = 1.32 \times 10^{-2}$ ). Interestingly, there were significant changes in the biosynthesis of steroids ( $P = 1.75 \times 10^{-4}$ ) between *cured* and naïve Huh7 cells. These results suggested that several lipid metabolism processes were substantially associated with efficient HCV replication in host cells.

Hierarchical clustering analyses of representative genes included in functional pathway categories

Based on pathway process analyses using the KEGG database, we performed hierarchical clustering analyses of each functional subset of genes (fold change  $> 1.2$ , Fig. 4a–c). The cell cycle, cholesterol biosynthesis and sphingolipid metabolism-related genes demonstrated clear clusters in replicon cells, *cured* cell and parental Huh7, respectively. In particular, cholesterol biosynthesis-related genes were activated in replicon cells and *cured* cells.

#### Mapping between pathway information and gene expression data

Knowing that cholesterol metabolism pathway was changed substantially in *cured* cells, we performed graphical mapping of the related genes to the KEGG Pathway map database (Fig. 5). Similar to the pathway analyses, cholesterol biosynthesis related genes, which are involved in the mevalonate pathway or sterol biosynthesis, were clearly activated in *cured* cells compared to naïve Huh7 (Fig. 5).

To verify the microarray results, we performed real-time RT-PCR of cholesterol biosynthesis-related genes including HMGR, SQLE, NSDHL, CYP51A1, TM7SF2 and EBP. All the genes were upregulated in replicon-expressing and *cured* cells compared to the naïve Huh7 cells (Fig. 6). These results were consistent with the microarray data.

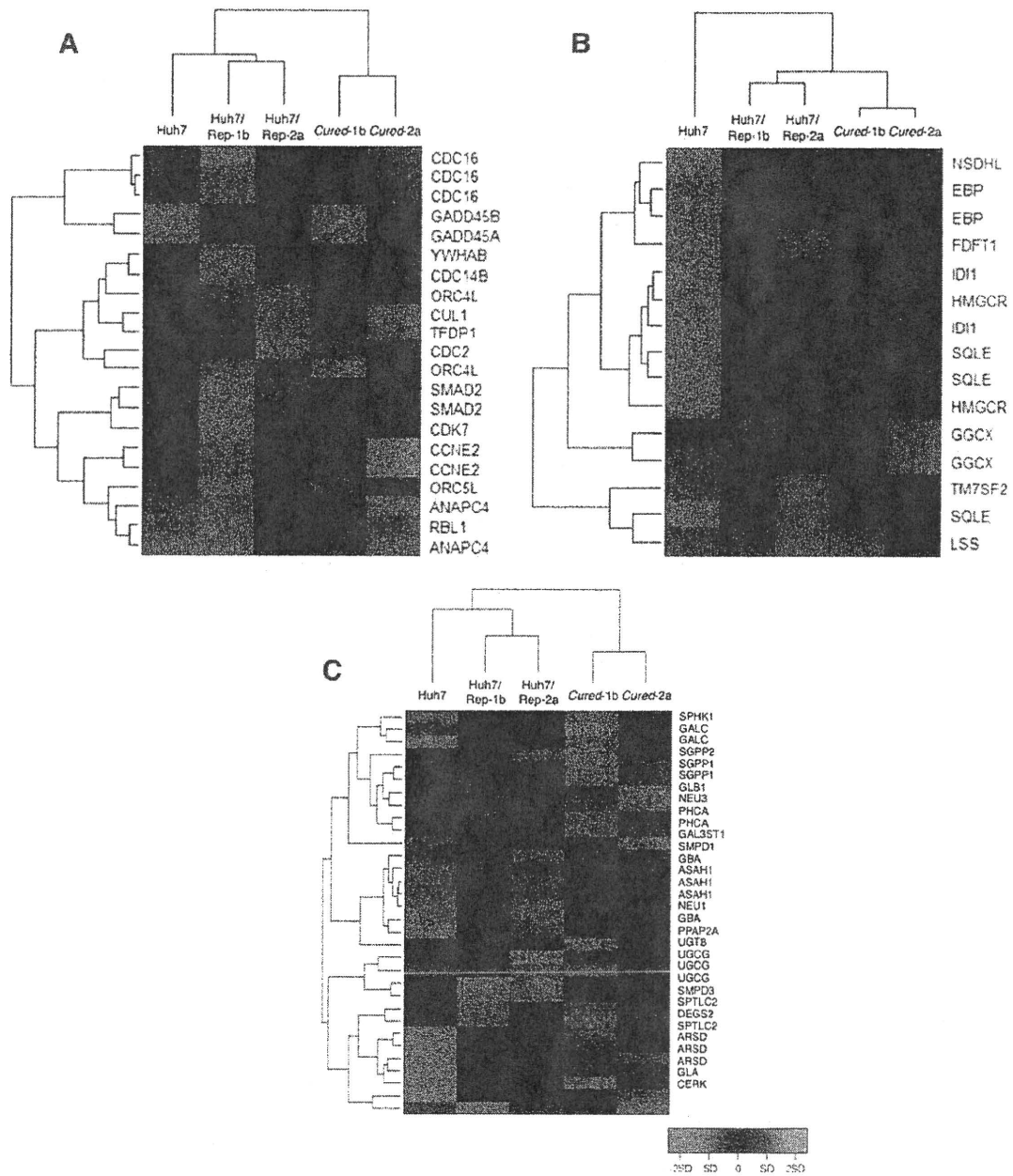
**Table 2** Microarray analysis: genes for which the expression changed more than two-fold in *cured-1b* and *cured-2a* cells compared to Huh7

Probe set	Title	Cured1b/Huh7		Cured2a/Huh7	
		Fold change	Q-value	Fold change	Q-value
210390_s_at	Chemokine (C–C motif) ligand 14/15	4.49	0.00	2.56	0.00
221168_at	PR domain containing 13	2.38	0.00	2.15	0.00
1553995_a_at	5 <sup>1</sup> -nucleotidase, ecto (CD73)	2.15	0.00	3.64	0.36
204897_at	Prostaglandin E receptor 4 (subtype EP4)	2.07	2.40	2.19	0.36
214522_x_at	Histone cluster 1, H2ad/H3d	3.01	4.12	2.98	0.46
214472_at	Histone cluster 1, H2ad/H3a-j	3.36	5.08	3.53	0.46
218280_x_at	Histone cluster 2, H2aa3/H2aa4	4.65	5.20	5.17	0.61
232035_at	Histone cluster 1, H4a-f, H4 h-l/histone cluster 2, H4a-b/histone cluster 4, H4	5.56	6.51	5.37	0.95
214455_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	4.72	6.51	2.68	0.61
202708_s_at	Histone cluster 2, H2be	4.48	6.51	5.31	0.95
214290_s_at	Histone cluster 2, H2aa3/H2aa4	4.06	6.51	2.98	2.08
209398_at	Histone cluster 1, H1c	3.40	6.51	3.60	2.08
230795_at	–	2.91	6.51	4.51	0.95
1553994_at	5 <sup>1</sup> -nucleotidase, ecto (CD73)	2.29	6.51	2.31	3.34
208180_s_at	Histone cluster 1, H4a-f, H4 h-l/histone cluster 2, H4a, H4b/histone cluster 4, H4	6.27	7.65	3.40	1.46
215779_s_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	3.53	7.65	5.86	1.46
206110_at	–	3.82	9.12	10.68	0.00
206535_at	Solute carrier family 2 (facilitated glucose transporter), member 2	3.58	9.12	2.84	5.02
213880_at	Leucine-rich repeat-containing G protein-coupled receptor 5	2.81	9.12	3.02	5.02
210387_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	3.24	10.68	2.28	11.66
203044_at	Chondroitin sulfate synthase 1	2.36	11.85	2.67	2.08
207102_at	Aldo-keto reductase family 1, member D1 (delta 4-3-ketosteroid-5-beta-reductase)	2.14	13.35	5.58	0.36
219596_at	THAP domain containing 10	2.20	16.73	3.32	2.08
217997_at	Pleckstrin homology-like domain, family A, member 1	2.15	23.48	3.01	3.34
217996_at	Pleckstrin homology-like domain, family A, member 1	2.12	23.48	2.21	11.66
217165_x_at	Metallothionein 1F	3.36	29.61	3.25	11.66
213629_x_at	Metallothionein 1F	3.13	29.61	3.70	11.66
210524_x_at	–	2.39	29.61	2.46	17.81
206143_at	Solute carrier family 26, member 3	2.35	29.61	2.49	17.81
212859_x_at	Metallothionein 1E	3.46	35.93	4.07	11.66
208581_x_at	Metallothionein 1X	3.31	35.93	3.52	17.81
204326_x_at	Metallothionein 1X	3.24	35.93	3.49	17.81
211456_x_at	Metallothionein 1 pseudogene 2	3.19	35.93	3.71	17.81
206461_x_at	Metallothionein 1H	3.09	35.93	3.49	17.81
216336_x_at	Metallothionein 1E, 1H, 1 M/metallothionein 1 pseudogene 2	3.02	35.93	3.27	17.81
212185_x_at	Metallothionein 2A	2.92	35.93	2.85	17.81
204745_x_at	Metallothionein 1G	2.73	35.93	3.31	17.81

#### Detection of intracellular lipid droplets in naïve, replicon-expressing and cured cells

Because several lipid-related pathways were extracted (Supplementary Tables 2 and 3), we examined phenotypes of the cell lines featuring different lipid metabolism

gene expression profiles by carrying out detection of cellular lipid droplets (Fig. 8a, b). The cells were stained by Oil red O or BODIPY493/503, dye solutions specific for neutral lipids. We found a large number of lipid droplets in the cytoplasm of each Huh7 cell line. The number of lipid droplets obviously was increased more in



**Fig. 4** Hierarchical clustering of representative genes included in each KEGG Pathway map. **a** Cell cycle, **b** cholesterol biosynthesis, **c** sphingolipid metabolism. Dendrograms shows the classification

determined by hierarchical clustering analysis. *Red and green colors* indicate relative overexpression and underexpression, respectively

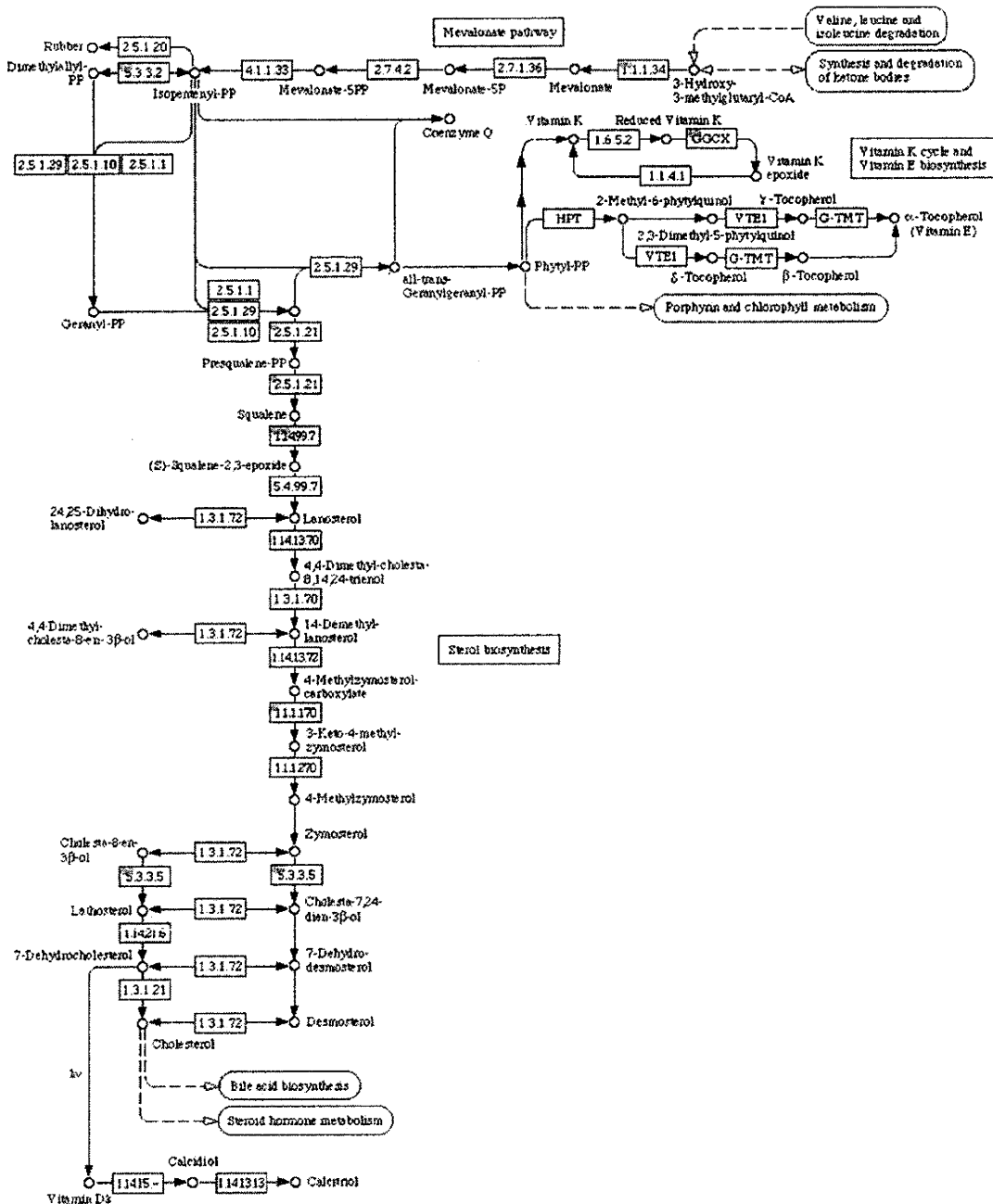
the two replicon-expressing cells and the *cured* cells than in the parental Huh7 cells. Lipid and HCV-NS5A double staining showed an increase in lipid droplets in cells that expressed HCV proteins (Fig. 8b). Analyses of the KEGG fatty acid metabolism pathway showed that a substantial number of the genes of these pathways were up-regulated in the *cured* cells compared to the naïve cells, although these could not reach statistical significance (Fig. 7).

**Effects of hepatitis C virus replication by PPAR-alpha and gamma agonists**

To assess the effects of lipid metabolic status on the intracellular replication of the HCV genome, Huh7/Rep-Feo cells were cultured with various concentrations of several PPAR-alpha agonists (clofibrate, fenofibrate and bezafibrate) and gamma agonists (pioglitazone and troglitazone) (Fig. 9). The luciferase activities of the Huh7/



BIOSYNTHESIS OF STEROIDS

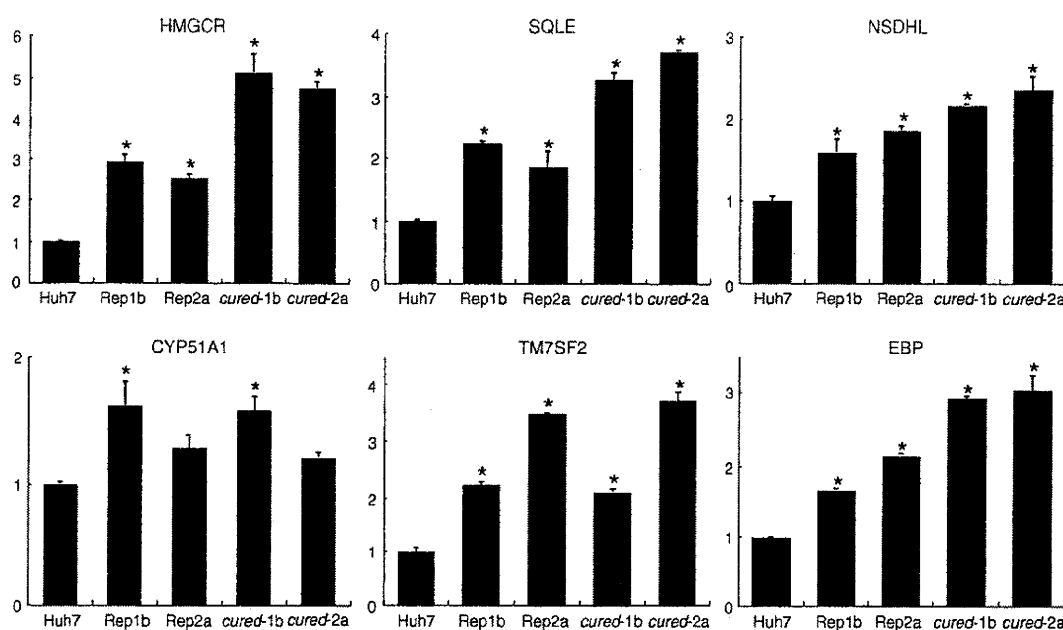


**Fig. 5** KEGG Pathway map and array data (biosynthesis of steroids). Gene expression changes were mapped on the pathways. Each circle within a box represents the corresponding probe set on Human Genome U133 Plus 2.0 array because multiple probe sets are sometimes designed for a single gene. Red circles indicate overexpressed genes in cured cells compared to parental Huh7 cells. The

dotted numerical code in each box represents the Enzyme Commission (EC) number based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). Correspondence between the genes that were examined in the microarray analyses and enzymes that are presented in Fig. 5 is shown in Supplementary Table 4

Rep-Feo cells showed that the replication of the HCV replicon was suppressed by clofibrate and fenofibrate in a dose-dependent manner, whereas pioglitazone and troglitazone elevated expression levels of replicon. The MTS

assay did not show any effect on cell viability or replication. These results suggest that the decrease or increase in HCV replication is due to specific effects of PPAR-alpha or gamma agonists on HCV replication.



**Fig. 6** Real-time detection RT-PCR. Real-time RT-PCR was performed to verify expression levels of genes that were listed in the cholesterol biosynthesis pathway in Fig. 4c and that showed

differences in their expression levels by microarray analyses. Assays were done in triplicate, and asterisks indicate *P*-values of less than 0.05

## Discussion

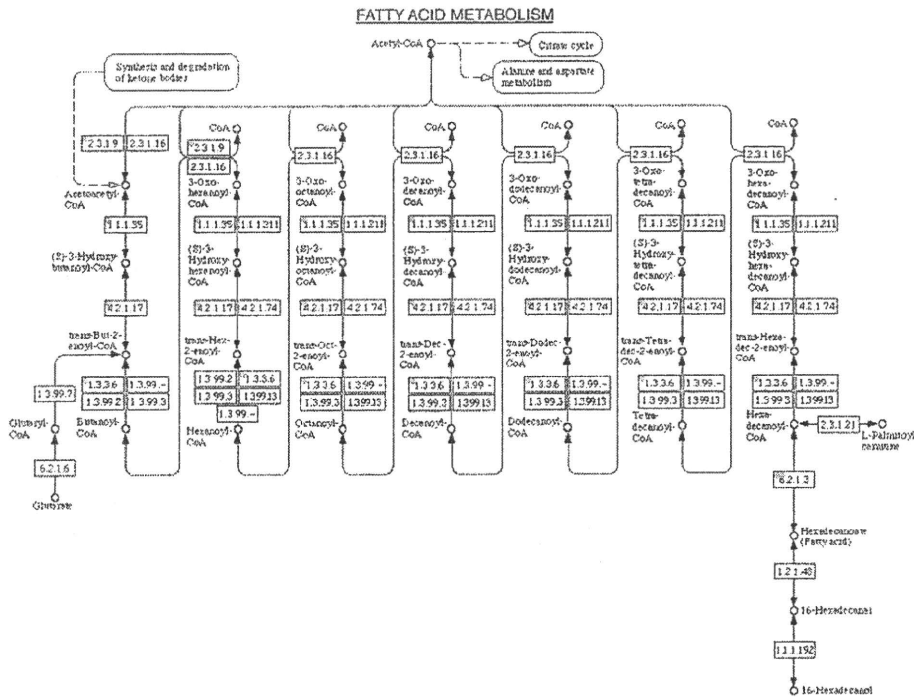
In our present analyses, we identified MAPK signaling, biosynthesis of steroid related and TGF- $\beta$  signaling pathways as significantly changed pathway processes by comparing replicon-expressing and *cured* cells (Supplementary Table 2). The results suggest that these pathways were primarily affected by HCV replication. Comparison of *cured* cells and naïve Huh7 cells identified cell cycle, TGF- $\beta$ , sphingolipid metabolism, and biosynthesis of steroids pathways as significantly changed pathways. Interestingly, cholesterol biosynthesis pathways were significantly changed in both comparisons (Supplementary Tables 2, 3). These data suggest that these pathways may positively regulate cellular HCV replication and that cholesterol biosynthesis pathways are primarily activated by HCV replication and may be essential for continuous virus replication.

There are several studies that report gene expression changes in replicon-expressing Huh7 cells as compared with the naïve cells [30–32]. In those studies, however, the changes in gene expression do not only reflect the effect of intracellular HCV replication, but also reflect alteration of host cell clonalities. Indeed, there are inconsistencies among studies. Use of the *cured* Huh7 cells can minimize the effect of cellular clonal changes because such Huh7 subclones have already been selected through HCV replicon transduction, drug-resistance selection and subsequent HCV elimination [33]. In our study, we have compared

gene expression between genotype 1b and 2a replicon cells, respective *cured* cells and the naïve parental cells, and have identified molecular signaling or metabolic pathways that were differentially up- or down-regulated over different HCV genotypes.

Comprehensive microarray analyses and pathway analyses were very useful for the identification of molecular mechanisms of HCV infection and replication in the host cells. We used the KEGG Pathway database [28], a knowledge-based database of biological systems that integrates genomic, chemical and systemic functional information. KEGG provides a reference knowledge base for linking genome to life through the process of PATHWAY mapping, which is to map, for example, a genomic or transcriptomic content of genes to KEGG reference pathways to infer systemic behavior of the cells or the organism. These pathway databases are free on-line resources. Using these analyses, the close relation between cholesterol metabolism and HCV replication was demonstrated. Moreover, in relation to this, when we examined the pathways of other lipid metabolism, it was shown that fatty acid biosynthesis metabolism-related pathways were significantly changed in *cured* cells, and indeed we found a large number of lipid droplets in the cytosol of replicon cells and *cured* cells.

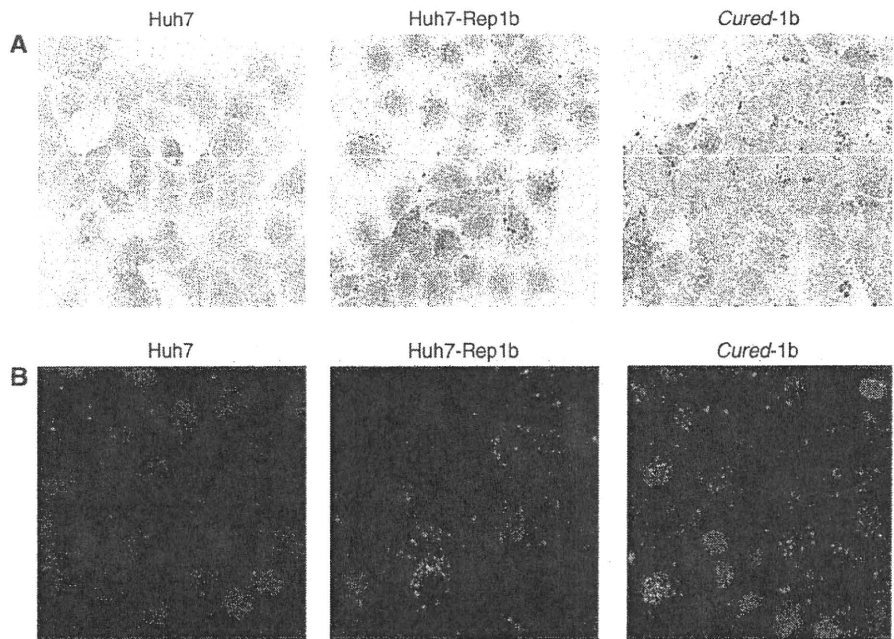
The HCV-JFH1 strain is the basis of a robustly replicating cell culture system reported recently [5]. We have performed comprehensive gene expression analyses using the HCV-JFH1 and the *cured* Huh7.5.1 cell line [6]. The



**Fig. 7** KEGG Pathway map and array data (fatty acid metabolism). Gene expression changes were mapped on the pathways. Each circle within a box represents the corresponding probe set on Human Genome U133 Plus 2.0 array because multiple probe sets are sometimes designed for a single gene. Red circles indicate overexpressed genes in

cured cells compared to parental Huh7 cells. The dotted numerical code in each box represents the Enzyme Commission (EC) number. Correspondence between the genes that were examined in the microarray analyses and enzymes that are presented in Fig. 7 are shown in Supplementary Table 4

**Fig. 8** Detection of intracellular lipid droplets and HCV NS protein. **a** Huh7 cells, replicon cells and cured cells were fixed and stained with Oil red O and Mayer's hematoxylin. Intracellular lipid droplets were detected as red spheres in the cells. Nuclei are stained in blue. **b** Rep1b/Huh7 cells were labeled with antibodies against NS5A (red). Lipid droplets and nuclei were stained with BODIPY493/503 (green) and DAPI (blue), respectively



KEGG Pathway analyses have identified several significantly affected pathways that are involved in the cell cycle, TGF-beta signaling, PPAR signaling and sterol

biosynthesis. These findings are consistent with our present results using the HCV subgenomic replicon (see the Supplementary Table 5; Supplementary Figs. 4, 5).

