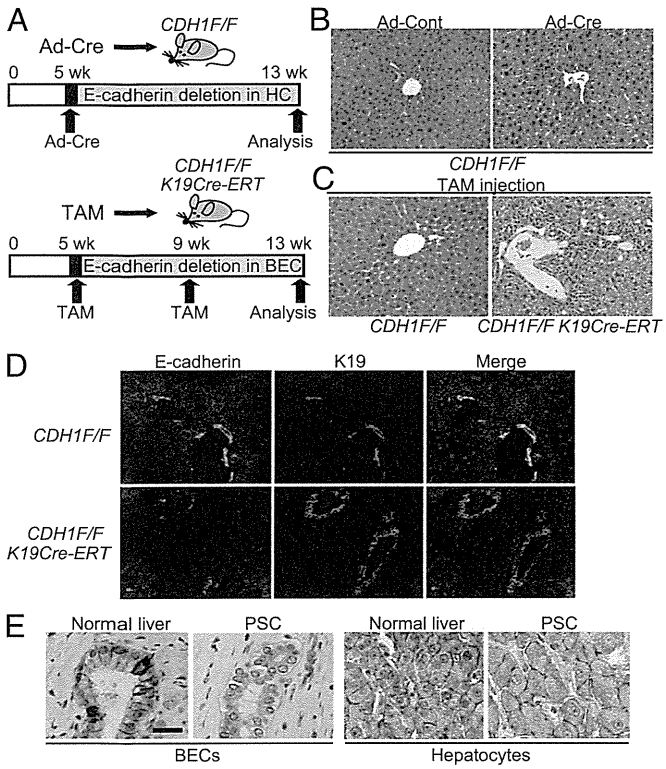

Fig. 2. Loss of E-cadherin in BECs rather than hepatocytes is a causal factor in periportal inflammation. (*A*) Experimental protocol of E-cadherin deletion in hepatocytes (HC) (*Upper*) or BECs (*Lower*). (*B*) H&E staining of liver sections from Ad-Cont– or Ad-Cre–injected $CDH1^{F/F}$ mice at 8 wk after adenovirus injection (×200). (*C* and *D*) Histological analyses of $CDH1^{F/F}$ and $CDH1^{F/F}/K19^{CreERT}$ mice at 8 wk after TAM injection. H&E staining (×200) (*C*) and double IF staining of E-cadherin (green) and K19 (red) (×400) (*D*). (*E*) IHC of E-cadherin in human normal liver and PSC samples ($n = 7$). (Scale bar, 20 μm.)

expression in hepatocytes was well-preserved in all samples. Based on these data, abnormality of E-cadherin expression in BECs might contribute to the pathogenesis of PSC.

Ductular Reaction in $CDH1^{\Delta L}$ Mice. To further characterize $CDH1^{\Delta L}$ mice, we conducted cDNA microarray analysis using whole-liver samples from young (2-mo-old) and aged (11-mo-old) $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mice. Interestingly, expression of progenitor cell markers such as Sox9, CD44, and Epcam, as well as inflammatory cytokines and chemokines, was up-regulated in both young and aged $CDH1^{\Delta L}$ mice compared with age-matched $CDH1^{F/F}$ mice (Fig. S3*A*). The microarray results were confirmed by IHC of these markers, and we found numerous primitive duct cells expressing progenitor cell markers, the so-called ductular reaction, in the periportal area (Fig. S3*B*). IHC of Ki67 revealed significant ongoing proliferation of these primitive duct cells but not hepatocytes in $CDH1^{\Delta L}$ mice (Fig. S3 *B* and *C*). Inflammatory cytokines have been implicated in hepatic progenitor cell induction (19), and $CDH1^{\Delta L}$ mice revealed infiltration of numerous F4/80-positive macrophages in the periportal area (Fig. S3*D*). Macrophage depletion by liposomal clodronate remarkably decreased the expression of interleukin 6, which is one of the major progenitor cells inducing cytokines (19), and subsequently reduced the ductular reaction (Fig. S3 *E–G*). Thus, the ductular reaction was due, in part, to macrophage-mediated inflammatory response.

Loss of E-Cadherin Accelerates Liver Tumorigenesis. A few of the male $CDH1^{\Delta L}$ mice (2/12, 16.7%) spontaneously developed liver tumors at 11 mo of age (Fig. S4 *A* and *B*). Most of the tumors were α-fetoprotein (AFP)–positive HCC, and negative for E-cadherin (Fig. S4 *C–E*). However, the tumor incidence rate in $CDH1^{\Delta L}$ mice was too low to investigate the detailed mechanism of hepatocarcinogenesis. Because Ras signaling is frequently activated in human HCC, we generated liver-specific active *Kras*-expressing $CDH1^{\Delta L}$ mice ($Kras/CDH1^{\Delta L}$) by crossing $CDH1^{\Delta L}$ mice with $Kras^{G12D}$ conditional knockin ($LSL\text{-}Kras^{G12D}$) mice. Strikingly, all male $Kras/CDH1^{\Delta L}$ mice developed multiple liver tumors at 8 mo of age ($n = 10$), whereas only 4 of 10 male $Alb\text{-}Cre/LSL\text{-}Kras^{G12D}/CDH1$ wild-type mice ($Kras/CDH1^{+/+}$) developed a few macroscopically visible small tumors at the same age (Fig. 3*A*). Quantitative analyses also showed significant acceleration of liver tumorigenesis by E-cadherin loss (Fig. 3*B*). We confirmed the absence of E-cadherin in both nontumor and tumor tissue in $Kras/CDH1^{\Delta L}$ mice (Fig. 3*C*). Most of the tumors arising in $Kras/CDH1^{\Delta L}$ mice were AFP-positive HCC and ranged from the typical trabecular type to a poorly differentiated type (Fig. 3*D*). Importantly, CCC and mixed-type HCC/CCC tumors were also seen in 10% and 40% of $Kras/CDH1^{\Delta L}$ mice, respectively (Fig. 3*D*). On the other hand, tumors in $Kras/CDH1^{+/+}$ mice were mostly AFP-negative dysplastic nodules or well-differentiated HCC (Fig. 3*D*). These results suggest that loss of E-cadherin cooperates with Ras signaling in liver cancer development.

Next, we assessed activation of extracellular signal-regulated kinase (ERK), which is a major downstream transducer of Ras, in nontumor tissue. ERK phosphorylation was almost undetectable by IHC in $CDH1^{F/F}$ mouse liver and faintly detected in periportal hepatocytes of $CDH1^{\Delta L}$ mice. $Kras/CDH1^{+/+}$ mice had diffuse ERK phosphorylation that was further increased in $Kras/CDH1^{\Delta L}$ mice (Fig. 3*E*). These findings were confirmed by Western blotting (WB) (Fig. S5*A*). To eliminate the influence of environmental factors on ERK activation, we isolated primary hepatocytes from $CDH1^{F/F}$, $LSL\text{-}Kras/CDH1^{+/+}$, and $LSL\text{-}KRas/CDH1^{F/F}$ mice (all Cre-negative), and then induced gene recombination by Ad-Cre infection. Ad-Cre efficiently induced gene recombination, and Ras activation and E-cadherin loss cooperatively increased ERK phosphorylation (Fig. 3*F*). This finding was confirmed in an immortalized human normal hepatocyte (20) by transfection of active Ras and knockdown of E-cadherin (Fig. S5*B*). E-cadherin was reported to inhibit the phosphorylation of several receptor tyrosine kinases (RTKs) (21), and phospho-RTK array analysis revealed that Ad-Cre–mediated E-cadherin deletion significantly increased only epithelial growth factor receptor (EGFR) phosphorylation (Fig. S5*C*). Thus, the cooperative activation of ERK by Ras and E-cadherin loss-induced EGFR phosphorylation may explain the accelerated carcinogenesis in $Kras/CDH1^{\Delta L}$ mice.

Furthermore, the ductular reaction was also significantly induced in $Kras/CDH1^{\Delta L}$ mice, likely due to increased ERK activation in the primitive duct cells (Fig. 3*G* and Fig. S5*D*). In addition to quantitative change, morphology of proliferating duct cells in $Kras/CDH1^{\Delta L}$ mice was sometimes dysplastic compared with that in $CDH1^{\Delta L}$ mice (Fig. 3*H*). These findings, along with CCC and mixed-type tumors, suggested that some tumors in $Kras/CDH1^{\Delta L}$ mice might originate from primitive duct cells, including bipotential progenitor cells induced by inflammation.

Spontaneous EMT, Vascular Invasion, and Intrahepatic Metastasis in $Kras/CDH1^{\Delta L}$ Tumors. EMT is considered a key process driving tumor invasiveness, and loss of E-cadherin expression is a hallmark of EMT (14). However, whether loss of E-cadherin is a consequence or a cause of EMT remains controversial. Importantly, in some tumors arising in $Kras/CDH1^{\Delta L}$ mice, HCC cells gradually transformed into fibroblast-like cells, and these cells are positive for both the HCC marker AFP and the mesenchymal marker vimentin, indicating that spontaneous EMT occurred in the tumors of these mice (Fig. 4 *A* and *B*). As expected, these cells were negative for E-cadherin (Fig. 4*B*). EMT was shown to be associated with a gain of stem cell properties (22). Indeed, evident expression of stem cell markers CD44 and Sox9 was seen in tumor cells undergoing EMT (Fig. 4*B*). We also found several tiny nodular lesions near the large

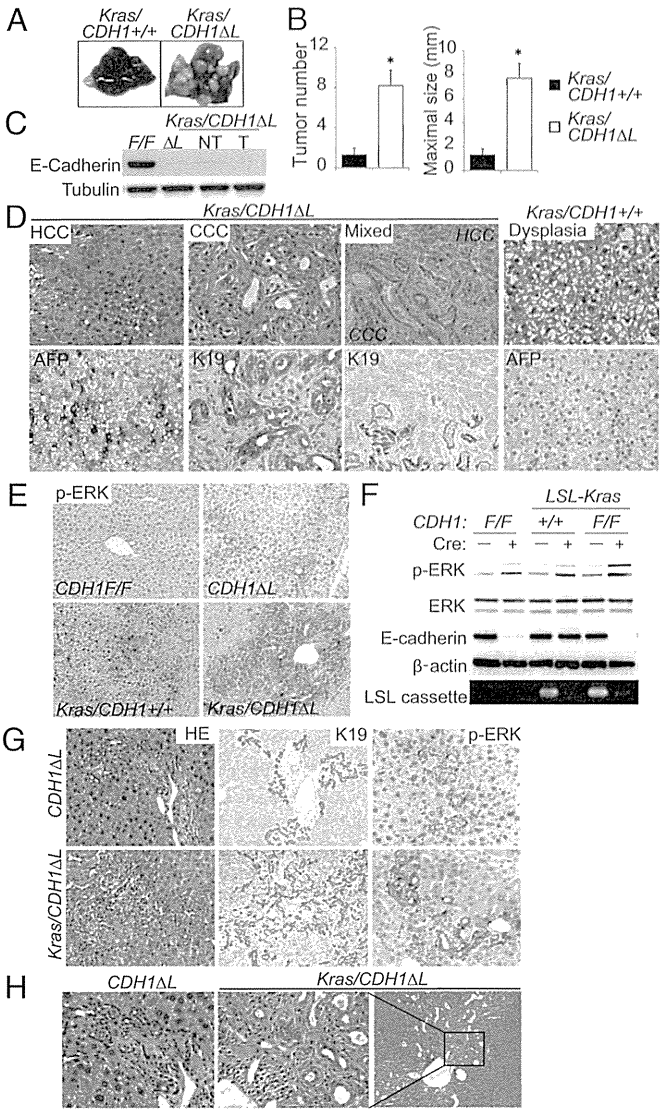

Fig. 3. Loss of E-cadherin cooperates with Ras activation in liver cancer development. (*A*) Representative images of the liver from 8-mo-old *Kras/CDH1*$^{+/+}$ and *Kras/CDH1*$^{\Delta L}$ mice. (*B*) Bar graphs of tumor number and maximal tumor size in each mouse are shown. Data are means ± SEM ($n = 10$ per group; $^{*}P < 0.05$). (*C*) WB analysis of E-cadherin expression in nontumor (NT) and tumor (T) tissues in *Kras/CDH1*$^{\Delta L}$ mice. (*D*) H&E staining and IHC of AFP and K19 of tumors from *Kras/CDH1*$^{\Delta L}$ and *Kras/CDH1*$^{+/+}$ mice (mixed, ×200; others, ×400). (*E*) IHC analysis of ERK phosphorylation in nontumor areas of 8-mo-old mouse livers (×200). (*F*) Primary hepatocytes isolated from *CDH1*$^{F/F}$, *LSL-Kras*, and *Kras/CDH1*$^{F/F}$ mice were infected with Ad-Cont or Ad-Cre. The indicated proteins were assessed by WB 48 h after infection. Cre-mediated recombination of the *LSL-Kras* allele (deletion of LSL cassette) was confirmed by PCR. (*G*) Ductular reaction assessment by H&E and IHC of K19 and ERK phosphorylation in primitive duct cells in 8-mo-old mice (H&E and K19, ×200; p-ERK, ×400). (*H*) H&E staining of the ductular reaction with dysplastic change in *Kras/CDH1*$^{\Delta L}$ mice (*Right*, ×100; *Left* and *Center*, ×400).

tumors undergoing EMT. Despite their small size, these lesions resembled the adjacent larger tumor in that they were strongly positive for vimentin, CD44, and Sox9, suggesting that they might represent intrahepatic metastasis (Fig. 4*C*). However, although tumor cells frequently invaded the central vein lumen in *Kras/CDH1*$^{\Delta L}$ mice (Fig. 4*D*), we could not find any apparent extrahepatic metastasis in any tissues, including the lymph nodes.

Loss of E-cadherin has been reported to induce β-catenin nuclear translocation that can promote carcinogenesis (23). Although the β-catenin expression was increased in the cancer cell membrane in *Kras/CDH1*$^{\Delta L}$ mice, we could not detect any nuclear translocation in either nontumor or tumor areas (Fig. S6*A*). In addition, E-cadherin knockdown in HCC cell lines did not induce significant β-catenin nuclear translocation (Fig. S6*B*). Thus, this mechanism does not play a major role in enhanced carcinogenesis by loss of E-cadherin in the liver.

Loss of E-Cadherin Promotes Chemically Induced HCC. To further examine the tumor-suppressive role of E-cadherin, *CDH1*$^{F/F}$ mice and *CDH1*$^{\Delta L}$ mice were injected with diethylnitrosamine (DEN) on postnatal day 14 (20, 24). After 8 mo, *CDH1*$^{\Delta L}$ mice had significantly increased number and size of liver tumors and developed histologically more advanced HCC compared with *CDH1*$^{F/F}$ mice (Fig. S7 *A* and *B*). As seen in *Kras/CDH1*$^{\Delta L}$ mice, strong ERK phosphorylation was seen in tumors of DEN-treated *CDH1*$^{\Delta L}$ mice (Fig. S7*C*). In addition, although an obvious fibroblast-like morphological change was not seen, some tumors in *CDH1*$^{\Delta L}$ mice strongly expressed CD44 and vimentin, whereas very few tumors in *CDH1*$^{F/F}$ mice expressed these markers (Fig. S7*D*). This confirmed that loss of E-cadherin enhances activation of ERK and expression of stem cell and EMT markers in a chemically induced HCC model.

Correlating E-Cadherin Loss with Mesenchymal and Stem Cell Markers in Human HCC. To investigate whether E-cadherin loss correlates with mesenchymal and stem cell markers in human HCC, we examined the expression of E-cadherin, CD44, and vimentin in human HCC cell lines. There were significant inverse correlations, particularly between E-cadherin and CD44 (Fig. 5*A*). Among these cell lines, we chose three that expressed E-cadherin, Hep3B, HuH7, and Alexander, and assessed the effect of E-cadherin knockdown with siRNA. After 6 d of E-cadherin knockdown, all three cell lines exhibited elevated expression of mesenchymal markers such as N-cadherin and vimentin and underwent morphological changes that resulted in an elongated mesenchymal-like appearance (Fig. 5 *B* and *C*). In addition, invasion capacity was significantly increased by E-cadherin knockdown (Fig. 5*D*), suggesting that loss of E-cadherin can be a causal factor of EMT and invasive phenotype of HCC.

Furthermore, the expression of stem cell markers such as CD44 and Sox9 were also increased upon E-cadherin knockdown in Hep3B cells (Fig. 5*B*), as seen in mouse HCC models. However, HuH7 and Alexander cells did not show any significant changes in stem cell markers upon E-cadherin knockdown (Fig. 5*B*). As shown in microarray data (Fig. S3*A*), various inflammatory cytokines and chemokines were up-regulated in *CDH1*$^{\Delta L}$ mouse livers in vivo. Interestingly, tumor necrosis factor α (TNF-α) stimulation significantly increased CD44 expression in E-cadherin–knocked down HuH7 cells but not in control cells (Fig. 5*E*). In E-cadherin–knocked down Alexander cells, not TNF-α alone but TNF-α in combination with transforming growth factor β (TGF-β) and hepatocyte growth factor (HGF) significantly increased CD44 expression (Fig. 5*E*). These findings suggest that loss of E-cadherin can up-regulate CD44 expression cooperatively with the inflammatory environment.

E-cadherin knockdown in Hep3B and HuH7 cells increased expression of Snail (Fig. 5*B*), which is an upstream transcriptional repressor of E-cadherin that plays important roles in EMT induction. Also in vivo, tumor cells undergoing EMT in *Kras/CDH1*$^{\Delta L}$ mice strongly expressed Snail (Fig. S8*A*). Additional knockdown of Snail in Hep3B cells suppressed up-regulation of mesenchymal and stem cell markers caused by E-cadherin knockdown (Fig. 5*F*), suggesting that Snail may link E-cadherin loss to EMT.

Finally, we investigated the expression of E-cadherin, CD44, and vimentin using a human HCC tissue array. Expression of E-cadherin tended to be lost in advanced-stage HCC (percentage of E-cadherin loss: 33.3% in stage I/II and 55.6% in stage III; $P = 0.088$). CD44 and vimentin expression exhibited significant inverse correlations with E-cadherin expression (Fig. 5*G*).

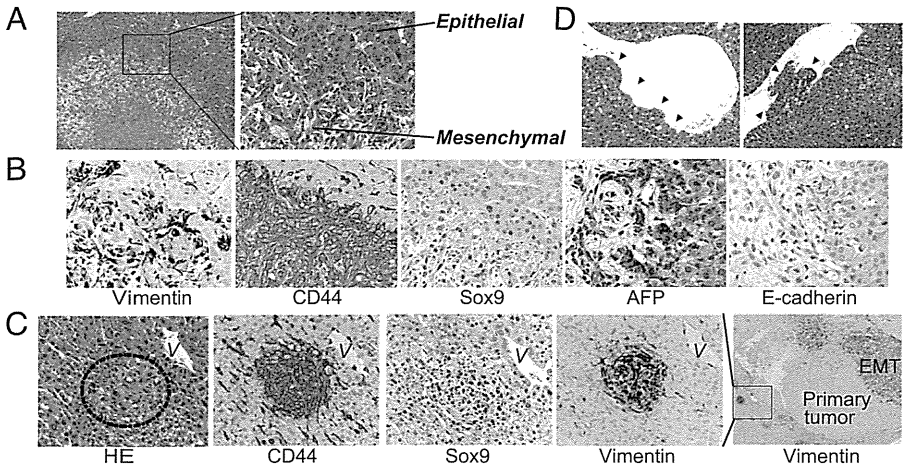

Fig. 4. Spontaneous epithelial–mesenchymal transition in tumors of *Kras/CDH1*$^{\Delta L}$ mice. (*A*) H&E staining of tumors from 8-mo-old *Kras/CDH1*$^{\Delta L}$ mice in which HCC cells gradually undergo an EMT-like morphological change (*Right*, ×400; *Left*, ×100). (*B*) Expression of the indicated proteins was assessed by IHC using serial sections (×200). (*C*) Intrahepatic metastasis-like lesions were seen near the large tumors undergoing EMT. Expression of EMT and stem cell markers in tiny nodular lesions was assessed by IHC using serial sections. V, vein near the lesion (*Right*, ×50; others, ×400). (*D*) H&E staining of tumor cell invasion into the central vein (×400). Arrowheads indicate invading cells.

Discussion

This study shows that E-cadherin plays critical roles in maintaining liver homeostasis. Although we assumed that periportal inflammation in *CDH1*$^{\Delta L}$ mice might be caused by disruption of the junction complex, electron microscopy showed no morphological abnormalities. This is consistent with a previous report (25). However, our functional analysis of the bile transport system using fluorescent-labeled bile acid showed impaired bile acid flow around zone 1. E-cadherin deletion in BECs by crossing *CDH1*$^{F/F}$ mice with *K19*^{CreERT} mice also triggered periportal inflammation similar to that seen in *CDH1*$^{\Delta L}$ mice. Therefore, loss of E-cadherin, especially in the intrahepatic bile duct epithelium, functionally impairs biliary flow, and subsequently induces cholestatic liver injury and sclerosing cholangitis. Our analysis of human PSC samples showed dysregulated E-cadherin expression in BECs, and another group also reported similar findings (26). Thus, abnormality of E-cadherin expression in BECs might contribute to the pathogenesis of PSC. Of note, PSC is well-known as an extraintestinal manifestation of ulcerative colitis (UC), and a recent genome-wide association study showed that single-nucleotide polymorphism in the CDH1 gene was associated with susceptibility to UC (27). Combined with our finding, impaired function of E-cadherin may be a common factor of these coexisting diseases.

Our data also suggest that E-cadherin is a tumor suppressor in the liver. A few *CDH1*$^{\Delta L}$ mice develop spontaneous HCC, and when knockout of E-cadherin is combined with Ras activation or chemical carcinogen administration, *CDH1*$^{\Delta L}$ mice display markedly accelerated carcinogenesis and an invasive phenotype. Although it has not been clear whether the loss of E-cadherin is a consequence or a cause of EMT, we have definitively demonstrated its causal role in vivo and in vitro. Recent reports have established a direct link between EMT and a gain of stem/progenitor cell properties (22), which is supported by our mouse models because tumor cells undergoing EMT clearly expressed stem cell markers. The expression of stem cell markers such as CD44 and Sox9 has been reported to be associated with poor prognosis in patients with HCC (28, 29). Interestingly, E-cadherin

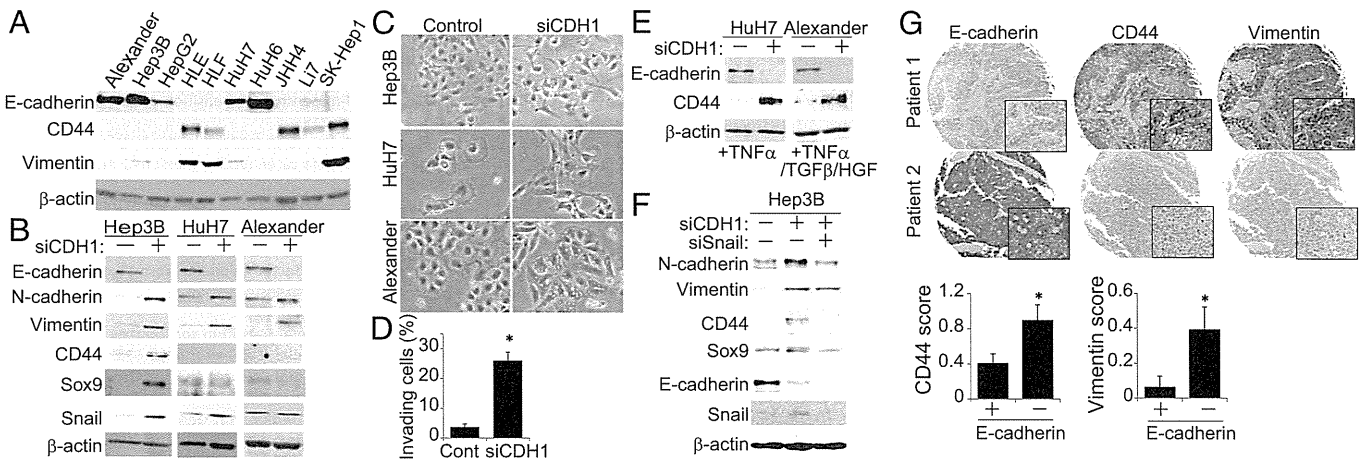

Fig. 5. Correlating E-cadherin loss with mesenchymal and stem cell markers in human HCC. (*A*) WB analysis of the indicated proteins in human HCC cell lines. (*B* and *C*) Effect of E-cadherin knockdown on human HCC cell lines. Expression levels of the indicated proteins (*B*) and morphological changes (*C*) in HCC cell lines 6 d after transfection with E-cadherin siRNA or controls. (*D*) Invasion capacity was assessed by invasion assay using Alexander cells. Bars show the percentages of invaded cells ($n = 3$). Data are means ± SD; $^*P < 0.05$. (*E*) The effect of E-cadherin knockdown on CD44 expression upon cytokine and/or growth factor stimulation in HuH7 and Alexander cells. Two days after transfection with E-cadherin or control siRNA, cells were stimulated with the indicated cytokines and/or growth factors (10 ng/mL TNF-α, 10 ng/mL HGF, or 1 ng/mL TGF-β). Four days later, the indicated proteins were assessed by WB. (*F*) Effect of E-cadherin and Snail double knockdown on the expression of mesenchymal and stem cell markers in Hep3B cells. Six days later, the indicated proteins were assessed by WB. (*G*) IHC of the expression pattern of E-cadherin, CD44, and vimentin using a human HCC tissue array that contains 60 HCC samples (E-cadherin–negative, $n = 28$; –positive, $n = 32$). When the number of E-cadherin–positive cells was <25% of the tumor cell population, the sample was defined as E-cadherin–negative. Representative images of two patients are shown (Magnifications: 40×; insets, 400×.). Data are means ± SD; $^*P < 0.05$.

knockdown was sufficient to up-regulate both EMT and stem cell markers in Hep3B cells, whereas additional supplementation of inflammatory cytokines and/or growth factors was necessary for up-regulation of CD44 in HuH7 and Alexander cells (Fig. 5*E*). Therefore, E-cadherin loss in the tumor cells and inflammatory environment might synergistically up-regulate stem cell markers in our mouse models.

E-cadherin interacts with various molecules and controls activation of related signaling pathways. In our study, increased ERK activation was implicated in accelerated tumorigenesis by loss of E-cadherin. It has been shown that E-cadherin can negatively regulate activation of divergent classes of RTKs such as EGFR, insulin-like growth factor 1 receptor, and c-Met by inhibiting receptor mobility (21); all of this can activate ERK signaling. Our analysis using the phospho-RTK array revealed that deletion of E-cadherin in primary hepatocytes increased EGFR phosphorylation without influencing other RTKs, suggesting that EGFR activation might be one explanation for E-cadherin loss-induced ERK phosphorylation in the liver.

The cellular source of liver cancers is a hot topic at present. *Kras/CDH1*$^{\Delta L}$ mice revealed not only an increase of the ductular reaction but also dysplastic change of ductal cells. These findings, along with CCC and mixed-type tumors, suggested that some tumors might originate from primitive duct cells, including bipotential progenitor cells. However, we also found some well-differentiated hepatocellular lesions that were adjacent to the central vein and distant from proliferating duct cells (Fig. S8*B*), suggesting that they may be derived from mature hepatocytes. In *Kras/CDH1*$^{+/+}$ mice, dysplastic foci were randomly found in zones 1, 2, and 3. Thus, we consider that the tumors in *Kras/CDH1*$^{\Delta L}$ mice may originate from both proliferating duct cells, including progenitor cells induced by inflammation and mature hepatocytes transformed by Ras activation. Meanwhile, recent studies suggested that mature hepatocytes also could transform to CCC (30, 31). Therefore, further analyses such as cell-lineage tracing are needed to definitively determine this issue.

In summary, loss of E-cadherin in the liver, especially in BECs, causes impairment of the intrahepatic biliary network and subsequent inflammatory reactions that lead to progenitor cell proliferation and periductal fibrosis. In some cases, these progenitor cells may lead directly to tumor development through oncogenic mutations such as Kras. In mature hepatocytes, loss of E-cadherin leads to ERK activation, EMT induction, and up-regulation of stem cell markers, which eventually results in enhanced carcinogenesis and an invasive phenotype (Fig. S9). Thus, E-cadherin plays critical roles in maintaining homeostasis and suppressing carcinogenesis in the liver.

Materials and Methods

CDH1$^{F/F}$, *Alb-Cre*, Lox-stop-lox Kras^{G12D}, and Rosa26-lox-stop-lox-YFP mice were purchased from The Jackson Laboratory. *K19*^{CreERT} mice were a generous gift from Guoqiang Gu (Vanderbilt University, Nashville, TN) (18). All mice were on the C57BL/6 genetic background. All animal experiments were approved by the Ethics Committee for Animal Experimentation at the Institute for Adult Diseases, Asahi Life Foundation, University of California, San Diego, and University of Tokyo, and conducted in accordance with the Guidelines for the Care and Use of Laboratory Animals. For details of reagents, biochemical analyses, histology, microarray, cell culture, and statistical analyses, please see *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Dr. Alan Hofmann for providing fluorescent-labeled bile acid and Dr. Shoen Kume (Kumamoto University) for handling the mice. This study was supported by a Grant-in-Aid for Scientific Research (25893042 to H.N.; 22300317 and 23114508 to S.M.; 23590933 to Y. Hirata), the Uehara Memorial Foundation (S.M.), and the Daiichi Sankyo Foundation of Life Science (H.N.). Work at the University of California, San Diego, was supported by the National Institutes of Health (CA118165 and CA155120).

1. Gumbiner BM (2005) Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol* 6(8):622–634.
2. Stepniak E, Radice GL, Vasioukhin V (2009) Adhesive and signaling functions of cadherins and catenins in vertebrate development. *Cold Spring Harb Perspect Biol* 1(5):a002949.
3. Larue L, Ohsugi M, Hirchenhain J, Kemler R (1994) E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc Natl Acad Sci USA* 91(17):8263–8267.
4. Tunggal JA, et al. (2005) E-cadherin is essential for in vivo epidermal barrier function by regulating tight junctions. *EMBO J* 24(6):1146–1156.
5. Boussadia O, Kutsch S, Hierholzer A, Delmas V, Kemler R (2002) E-cadherin is a survival factor for the lactating mouse mammary gland. *Mech Dev* 115(1-2):53–62.
6. Ihara A, Koizumi H, Hashizume R, Uchikoshi T (1996) Expression of epithelial cadherin and alpha- and beta-catenins in nontumoral livers and hepatocellular carcinomas. *Hepatology* 23(6):1441–1447.
7. Berx G, van Roy F (2009) Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb Perspect Biol* 1(6):a003129.
8. Matsumura T, Makino R, Mitamura K (2001) Frequent down-regulation of E-cadherin by genetic and epigenetic changes in the malignant progression of hepatocellular carcinomas. *Clin Cancer Res* 7(3):594–599.
9. Lim SO, et al. (2008) Epigenetic changes induced by reactive oxygen species in hepatocellular carcinoma: Methylation of the E-cadherin promoter. *Gastroenterology* 135(6):2128–2140, e8.
10. Lee S, Kim WH, Jung HY, Yang MH, Kang GH (2002) Aberrant CpG island methylation of multiple genes in intrahepatic cholangiocarcinoma. *Am J Pathol* 161(3):1015–1022.
11. Wei Y, et al. (2002) Altered expression of E-cadherin in hepatocellular carcinoma: Correlations with genetic alterations, beta-catenin expression, and clinical features. *Hepatology* 36(3):692–701.
12. Yang MH, et al. (2009) Comprehensive analysis of the independent effect of twist and snail in promoting metastasis of hepatocellular carcinoma. *Hepatology* 50(5): 1464–1474.
13. Ma L, et al. (2010) miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol* 12(3):247–256.
14. Lee JM, Dedhar S, Kalluri R, Thompson EW (2006) The epithelial-mesenchymal transition: New insights in signaling, development, and disease. *J Cell Biol* 172(7):973–981.
15. Osada T, et al. (1996) E-cadherin is involved in the intrahepatic metastasis of hepatocellular carcinoma. *Hepatology* 24(6):1460–1467.
16. Malato Y, et al. (2011) Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J Clin Invest* 121(12):4850–4860.
17. Akagi K, et al. (1997) Cre-mediated somatic site-specific recombination in mice. *Nucleic Acids Res* 25(9):1766–1773.
18. Means AL, Xu Y, Zhao A, Ray KC, Gu G (2008) A CK19(CreERT) knockin mouse line allows for conditional DNA recombination in epithelial cells in multiple endodermal organs. *Genesis* 46(6):318–323.
19. Duncan AW, Dorrell C, Grompe M (2009) Stem cells and liver regeneration. *Gastroenterology* 137(2):466–481.
20. Nakagawa H, et al. (2011) Apoptosis signal-regulating kinase 1 inhibits hepatocarcinogenesis by controlling the tumor-suppressing function of stress-activated mitogen-activated protein kinase. *Hepatology* 54(1):185–195.
21. Qian X, Karpova T, Sheppard AM, McNally J, Lowy DR (2004) E-cadherin-mediated adhesion inhibits ligand-dependent activation of diverse receptor tyrosine kinases. *EMBO J* 23(8):1739–1748.
22. Mani SA, et al. (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133(4):704–715.
23. Orsulic S, Huber O, Aberle H, Arnold S, Kemler R (1999) E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. *J Cell Sci* 112(Pt 8):1237–1245.
24. Maeda S, Kamata H, Luo JL, Leffert H, Karin M (2005) IKKbeta couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* 121(7):977–990.
25. Battle MA, et al. (2006) Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. *Proc Natl Acad Sci USA* 103(22):8419–8424.
26. Rygiel KA, et al. (2008) Epithelial-mesenchymal transition contributes to portal tract fibrogenesis during human chronic liver disease. *Lab Invest* 88(2):112–123.
27. Barrett JC, et al.; UK IBD Genetics Consortium; Wellcome Trust Case Control Consortium 2 (2009) Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet* 41(12):1330–1334.
28. Yang GH, et al. (2008) Osteopontin combined with CD44, a novel prognostic biomarker for patients with hepatocellular carcinoma undergoing curative resection. *Oncologist* 13(11):1155–1165.
29. Guo X, et al. (2012) Expression features of SOX9 associate with tumor progression and poor prognosis of hepatocellular carcinoma. *Diagn Pathol* 7:44.
30. Fan B, et al. (2012) Cholangiocarcinomas can originate from hepatocytes in mice. *J Clin Invest* 122(8):2911–2915.
31. Sekiya S, Suzuki A (2012) Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. *J Clin Invest* 122(11):3914–3918.

Supporting Information

Nakagawa et al. 10.1073/pnas.1322731111

SI Materials and Methods

Reagents. The following antibodies were used in these experiments: anti–E-cadherin, anti–phospho-ERK, anti-vimentin, and anti-CD44 for Western blotting (Cell Signaling); anti-CD44 for immunohistochemistry and anti-CD44v6 (AbD Serotec); anti-Sox9 and anti-K19 (Santa Cruz Biotechnology); anti–N-cadherin (Millipore); anti-actin and anti-tubulin (Sigma); anti-EpCAM (Abcam); anti–α-smooth muscle actin (Dako); anti-CD45 (eBioscience); anti-F4/80 (Caltag); anti-Ki67 (GeneTex); anti–α-fetoprotein (AFP) (Biocare Medical); anti–β-catenin (GeneTex and BD Transduction Laboratories); and anti-Snail (Novus Biologicals). Recombinant human TNF-α, hepatocyte growth factor, and TGF-β were purchased from R&D Systems.

Immunoblotting. Whole-liver protein homogenates were subjected to SDS/PAGE and transferred to a polyvinylidene membrane (Amersham Biosciences). The membrane was probed with primary antibodies and then incubated with the secondary antibody. Immunocomplexes were detected using the ECL System (Amersham Biosciences). Phospho-RTK (receptor tyrosine kinase) array was performed using the Proteome Profiler Mouse Phospho-RTK Array Kit (R&D Systems) according to the manufacturer's protocol.

RNA Extraction and Real-Time PCR. RNA was extracted from liver tissue using ISOGEN (Nippon Gene). First-strand cDNA was synthesized using SuperScript 2 (Invitrogen Life Technologies), and the relative amount of each mRNA was quantified by real-time PCR and normalized against GAPDH mRNA expression. Primer sequences are available upon request.

Histology. Livers were fixed in 10% (vol/vol) neutral-buffered formalin or 4% (wt/vol) paraformaldehyde, embedded in paraffin, sectioned, and stained with hematoxylin and eosin (H&E) or Sirius red, and processed for immunohistochemistry. In the case of immunostaining of E-cadherin, perfusion fixation was performed with 4% paraformaldehyde. For frozen block preparation, liver tissue was embedded in Tissue-Tek OCT compound (Sakura Finetek). For immunohistochemistry, fixed and paraffin-embedded liver sections were deparaffinized and incubated in citrate buffer at 95 °C for 40 min for antigen retrieval, and then incubated overnight at 4 °C with the primary antibodies. Biotinylated secondary antibodies (Pharmingen) were added and incubated for 20 min at room temperature. Streptavidin-horseradish peroxidase (Pharmingen) was added, and after 30 min the sections were developed with 3,3′-diaminobenzidine substrate and counterstained with hematoxylin. Quantification of the stained area was performed with ImageJ software (National Institutes of Health). Frozen slides were incubated with primary antibodies, followed by secondary antibodies labeled with Alexa 488 or 594 (Molecular Probes).

The expression of E-cadherin in human primary sclerosing cholangitis (PSC) samples was analyzed using liver explants and biopsy samples containing medium or large bile ducts. Adjacent nontumor tissue of metastatic liver cancer was used as a normal control. Liver explants were obtained through the Cooperative Human Tissue Network, and biopsy samples for the diagnosis of PSC were obtained at the University of Tokyo with approval of the medical ethics committee of the University of Tokyo and with informed consent.

The tissue array was purchased from US Biomax and contained 60 paired human hepatocellular carcinoma (HCC) and nontumor tissues. When the number of E-cadherin–positive cells was <25% of the tumor cell population, the sample was defined as E-cadherin–negative. CD44 and vimentin expression in the tumors were scored semiquantitatively based on a scale staining pattern of 0–3 (0, all tumor cells are negative; 1, ≤25%; 2, 25–50%; 3, >50% of the tumor cell population is positive for CD44 or vimentin).

TUNEL staining was performed using an ApoAlert DNA Fragmentation Assay Kit (Clontech).

Electron Microscopy. Livers were fixed with 2% glutaraldehyde/0.1 M phosphate buffer (pH 7.4) and then postfixed with 1% osmium tetroxide/0.1 M phosphate buffer (pH 7.4). Fixed samples were embedded in epoxy resin. Ultrathin sections were counterstained with uranyl acetate and observed with a JEOL 1200 EXII electron microscope.

Immunofluorescence Microscopy for Fluorescent-Labeled Bile Acid. Fluorescent-labeled bile acid was kindly provided by Alan Hofmann (University of California, San Diego), and 100 μL of a 2 mM solution was injected into the tail veins of 2-mo-old *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice. Fifteen minutes after injection, the mice were killed and the removed livers were embedded in OCT compound. Sections from frozen tissues were visualized by fluorescence microscopy.

Adenovirus- and Tamoxifen-Induced Gene Recombination in Vivo. Recombinant adenovirus expressing Cre-recombinase and control empty adenovirus vector were diluted in PBS and i.v. injected into mice (1×10^9 plaque-forming units per mouse). Tamoxifen (TAM) was dissolved in corn oil and intraperitoneally injected (4 mg TAM per mouse).

Microarray Analysis. Microarray analysis was performed using SurePrint G3 Mouse Gene Expression 8 × 60K arrays (Agilent Technologies) according to the manufacturer's protocol. Data were preprocessed using Agilent GeneSpring GX11.

Diethylnitrosamine-Induced Hepatocellular Carcinoma. Diethylnitrosamine (DEN) (Sigma) was dissolved in PBS and injected intraperitoneally into mice (25 mg/kg) on postnatal day 14. Mice were killed after 8 mo, and their livers were removed and examined for visible tumors.

Cell Culture, Transfections, and RNA Interference. HuH7, Alexander, HLF, HepG2, and SK-Hep1 cell lines as well as a human normal hepatocyte cell line were cultured in Dulbecco's modified Eagle medium with 10% FBS. Hep3B cells were cultured in minimum essential medium (MEM) with 2 mM L-glutamine, 0.15% sodium bicarbonate, 0.1 mM nonessential amino acid solution, 1 mM sodium pyruvate, and 10% FBS. JHH4 cells were cultured in MEMa with 10% FBS. Li7 cells were cultured in RPMI-1640 with 10% FBS. Mouse primary hepatocytes were isolated by the collagenase perfusion method and then cultured in William's E medium with 10% FBS on collagen-coated plates.

A small interfering RNA (siRNA) construct was obtained with siGENOME SMARTpool reagents (Dharmacon), and siRNA transfections were performed using RNAiMAX (Invitrogen). pCMV3 myc-tagged RasG12V plasmid was transfected into normal human hepatocyte cell lines using X-tremeGENE HP DNA Transfection Reagent (Roche). Subcellular protein fractionation was performed using the ProteoExtract Subcellular Proteome Extraction Kit (Calbiochem).

Invasion Assay. An invasion assay was performed using a BD BioCoat Matrigel invasion chamber containing 8.0-μm pore size polyethylene terephthalate membrane inserts in a 24-well format, according to the manufacturer's protocol. Briefly, 5 d after transfection with E-cadherin or control siRNA, 1×10^4 cells were added to the upper chambers in serum-free culture media. The lower chamber

contained 10% FBS. After 24 h, cells on the upper surface of each membrane were removed with cotton swabs, and those that successfully migrated to the lower surface were stained with the Diff-Quick Kit (Sysmex, Japan) and counted. Data are expressed as the percent invasion through the Matrigel and membrane relative to the number of cells migrating through the control membrane.

1. Grivennikov SI, et al. (2012) Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* 491(7423):254–258.

Mice. $APC^{F/F}$ mice have been described previously (1).

Statistical Analyses. Statistical analyses were performed using the Student *t* test or one-way analysis of variance followed by the Tukey–Kramer test for multiple comparisons. A P value <0.05 indicated statistical significance.

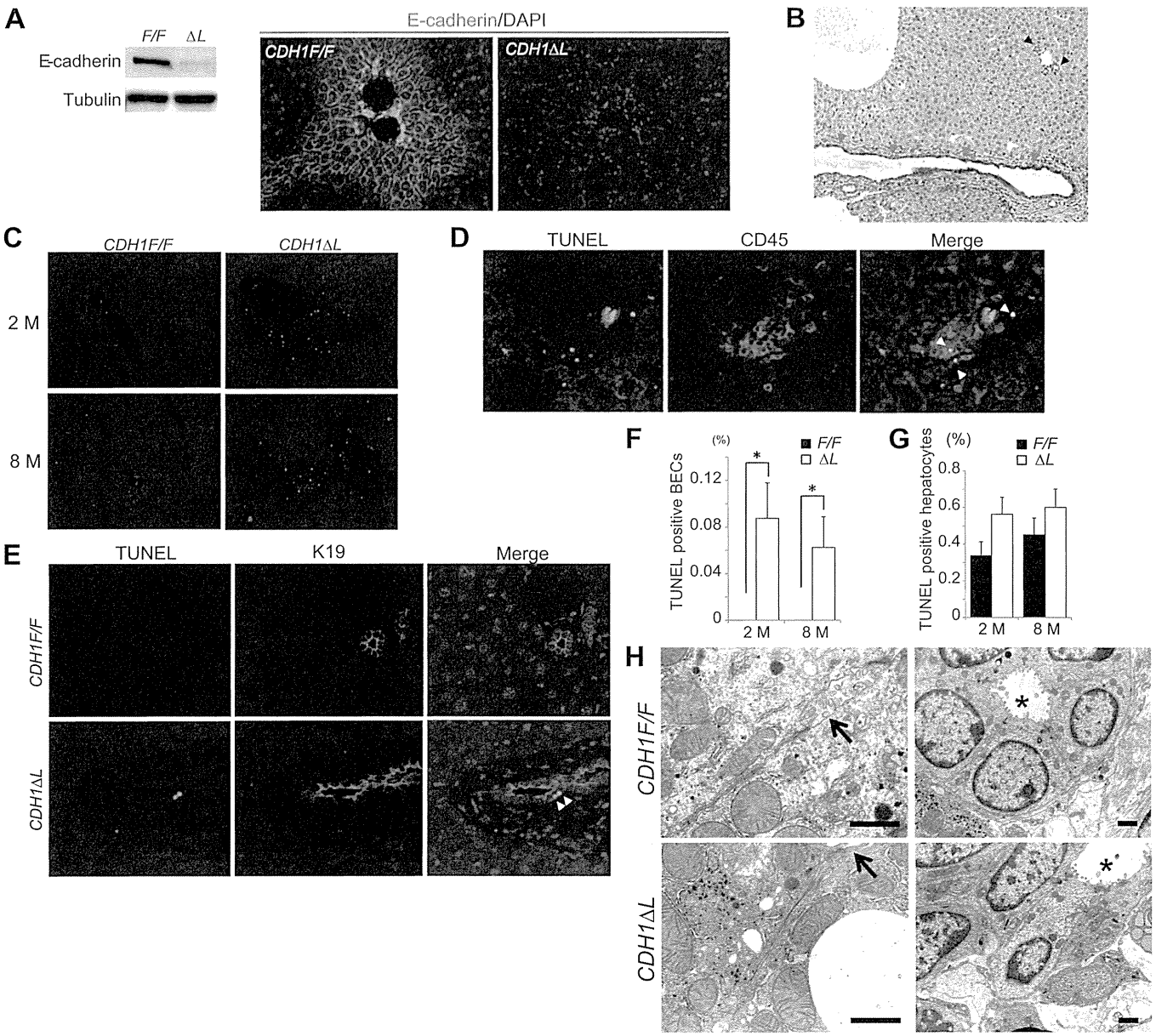

Fig. S1. Characterization of $CDH1^{\Delta L}$ mouse livers. (*A*) Analysis of E-cadherin expression by Western blotting of total liver protein and immunofluorescence (IF) staining of liver sections obtained from 1-mo-old $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mice (×200). Nuclei were labeled with DAPI (blue). (*B*) Immunohistochemistry (IHC) analysis of E-cadherin expression in 1-mo-old $CDH1^{\Delta L}$ mouse liver. Expression of E-cadherin was absent in the interlobular bile duct (black arrowheads) but not in the large bile duct near the common bile duct (red arrowheads) (×200). (*C*) Representative TUNEL-stained sections in 2- and 8-mo-old $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mouse livers (×200). (*D*) Double IF staining of TUNEL (green) and pan-leukocyte marker CD45 (red) in the periportal area of 2-mo-old $CDH1^{\Delta L}$ mouse liver (×400). Arrowheads indicate TUNEL-CD45 double-positive cells. (*E*) Double IF staining of TUNEL (green) and biliary epithelial cell (BEC) marker K19 (red) in 2-mo-old $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mouse livers (×400). Arrowheads indicate TUNEL-K19 double-positive cells. (*F* and *G*) Frequencies of TUNEL-positive duct cells (*F*) and hepatocytes (*G*) in 2- and 8-mo-old $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mouse livers. Data are expressed as means ± SEM ($n = 8$ per group; $^*P < 0.05$.). (*H*) Electron microscopic images of junction complexes in hepatocytes (*Left*) and bile duct cells (*Right*) in 2-mo-old $CDH1^{F/F}$ and $CDH1^{\Delta L}$ mice. (Scale bars, 1 μm.) Arrows and asterisks indicate bile canaliculi and bile duct lumens, respectively.

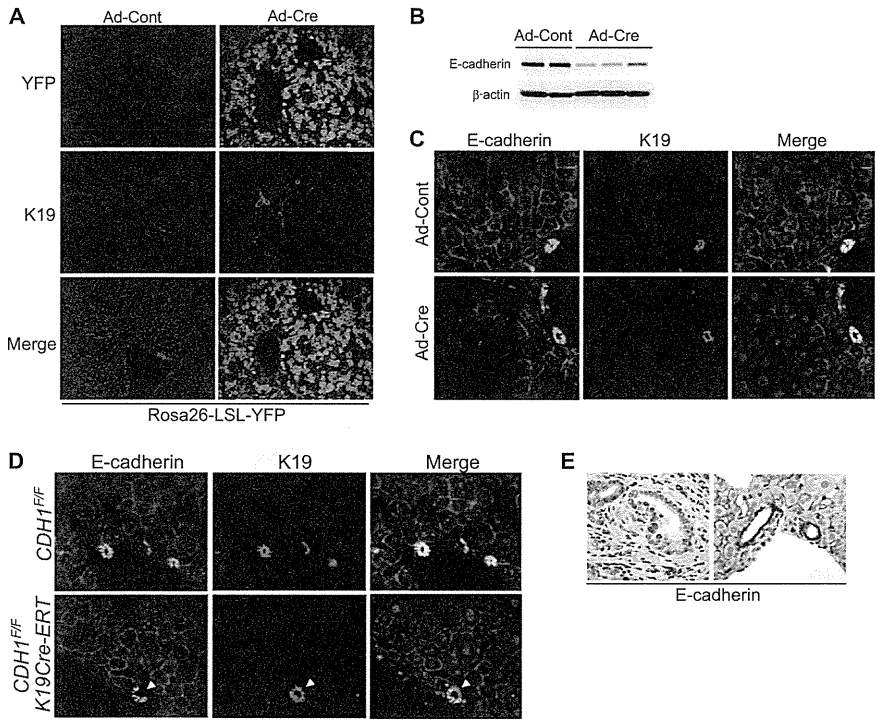

Fig. S2. Separated analysis of the role of E-cadherin in hepatocytes and BECs to maintain liver homeostasis. (*A*) Ad-Cre or Ad-Cont (control adenovirus) was i.v. injected into Rosa26-lox-stop-lox-YFP mice, and 1 wk later, hepatocyte-specific Cre-loxP recombination was confirmed by double IF staining of YFP (green) and K19 (red) (×200). (*B* and *C*) Analysis of E-cadherin expression in *CDH1*$^{F/F}$ mouse liver at 8 wk after Ad-Cre or Ad-Cont injection. Western blotting of total liver protein shows significant reduction of E-cadherin protein in Ad-Cre–injected mice compared with Ad-Cont–injected mice (*B*). Double IF staining of E-cadherin (green) and K19 (red) shows that expression of E-cadherin in hepatocytes was frequently deleted, whereas that in BECs was well-preserved (×400) (*C*). (*D*) TAM was intraperitoneally injected into *CDH1*$^{F/F}$ and *CDH1*$^{F/F}$/*K19*^{CreERT} mice, and 1 wk later, BEC-specific deletion of E-cadherin was confirmed by double IF staining of E-cadherin (green) and K19 (red) (×400). Arrowheads indicate E-cadherin–deleted K19-positive BECs. (*E*) IHC analysis of E-cadherin expression in *CDH1*$^{F/F}$/*K19*^{CreERT} mouse liver at 8 wk after TAM injection (×400). (*Left*) Efficient deletion of E-cadherin in BECs with strong periductal inflammation. (*Right*) Inefficient deletion without inflammation. Arrowheads indicate efficiently E-cadherin–deleted bile duct.

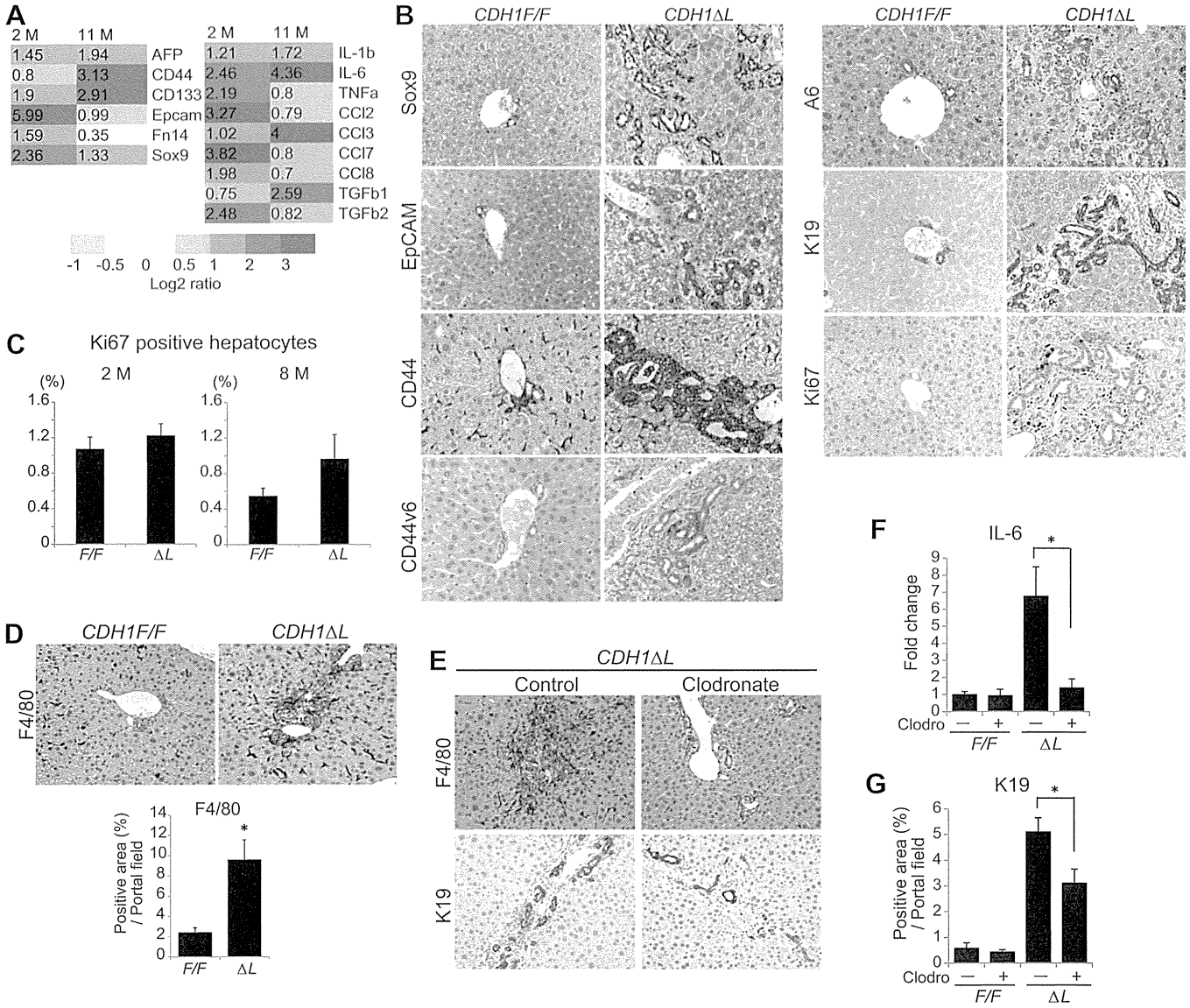

Fig. S3. Ductular reaction in *CDH1*$^{\Delta L}$ mice. (*A*) cDNA microarray analysis of whole-liver samples from *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice at 2 and 11 mo of age. Expression of stem cell markers (*Left*) and inflammatory cytokines and chemokines (*Right*) is up-regulated in *CDH1*$^{\Delta L}$ mice compared with *CDH1*$^{F/F}$ mice. Representative genes are shown and data are expressed as log2 ratio. (*B*) IHC analysis of the indicated stem cell and proliferation markers in 8-mo-old *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice (×400). (*C*) Frequencies of Ki67-positive hepatocytes in 2- and 8-mo-old *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mouse livers. Data are expressed as means ± SEM (n = 4–6 per group). (*D*) IHC analysis of F4/80 in 8-mo-old *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice (×200). The bar graph shows F4/80-positive areas. Data are expressed as means ± SD (n = 5 per group; $^*P < 0.05$). (*E–G*) Effect of macrophage depletion on the ductular reaction. Two-month-old *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice were i.v. injected with 100 μL liposomal clodronate or PBS liposome (control) four times every 4 d (days 1, 5, 9, and 13) (n = 3–4 per group). The mice were killed 4 d after the final injection. IHC analysis of the indicated proteins in liposomal clodronate or control-injected *CDH1*$^{\Delta L}$ mice (×200) (*E*). The relative level of IL-6 mRNA was determined by real-time PCR (*F*). K19-positive areas were quantified (*G*). Data are expressed as means ± SD; $^*P < 0.05$.

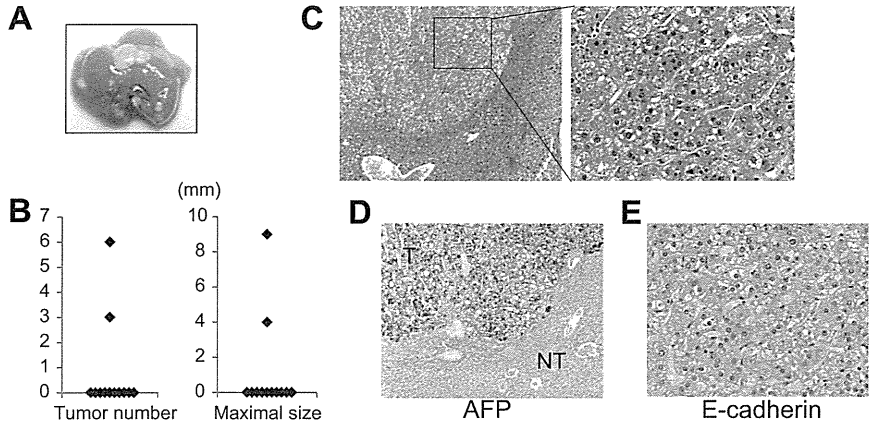

Fig. S4. Some *CDH1*$^{\Delta L}$ mice spontaneously develop liver tumors. (*A*) Representative image of a tumor-bearing liver in an 11-mo-old *CDH1*$^{\Delta L}$ mouse. (*B*) Graphs show tumor number and maximal tumor size in 11-mo-old *CDH1*$^{\Delta L}$ mice ($n = 12$). (*C*) H&E staining of a liver tumor in a *CDH1*$^{\Delta L}$ mouse (*Left*, ×100; *Right*, ×400). (*D* and *E*) IHC analysis of AFP and E-cadherin expression in tumor tissue (*D*, ×100; *E*, ×400). NT, nontumor; T, tumor.

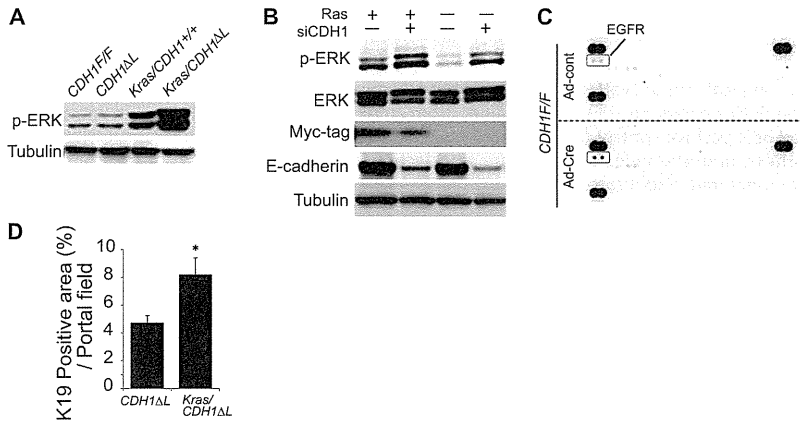

Fig. S5. Cooperative activation of ERK by active Ras and loss of E-cadherin. (*A*) Western blot evaluation of ERK phosphorylation in nontumor tissues of 8-mo-old mouse livers. (*B*) Normal immortalized human hepatocytes were transfected with myc-tagged active Ras (G12V) and siRNA targeted to E-cadherin. The indicated proteins were assessed by Western blotting. (*C*) Primary hepatocytes isolated from *CDH1*$^{F/F}$ mice were infected with Ad-Cont or Ad-Cre and, 48 h afterward, RTK phosphorylation was assessed by phospho-RTK array. (*D*) Quantification of a K19-positive area in 8-mo-old *CDH1*$^{\Delta L}$ and *Kras/CDH1*$^{\Delta L}$ mouse livers (related to Fig. 3*G*). The bar graph shows K19-positive areas. Data are expressed as means ± SD ($n = 5$ per group; $^*P < 0.05$).

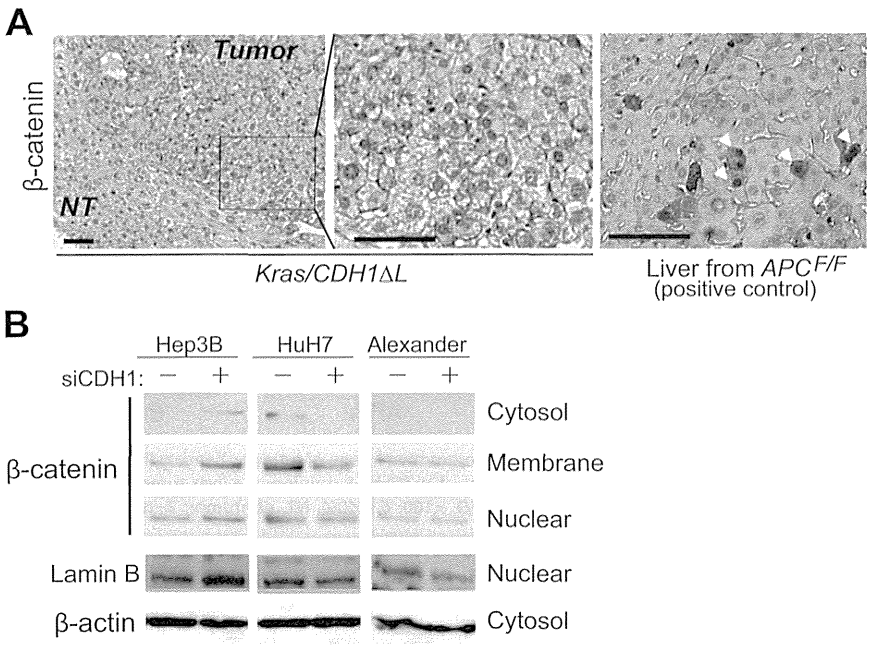

Fig. S6. Effect of E-cadherin loss on β-catenin activation. (*A*) Immunohistochemical analysis of β-catenin expression in tumor and nontumor areas of *Kras/CDH1*$^{\Delta L}$ mouse liver (*Left* and *Center*). (Scale bars, 50 μm.) Liver obtained from an APC flox/flox (*APC*$^{F/F}$) mouse, in which β-catenin/TCF signaling is activated due to reduced APC protein, was used as a positive control (*Right*). Arrowheads indicate hepatocytes positive for nuclear β-catenin. (*B*) Analysis of β-catenin expression in fractionated Hep3B, HuH7, and Alexander cells 6 d after transfection with E-cadherin siRNA or control siRNA. Proteins from the cytosolic, membrane, and nuclear fractions were separated, and β-catenin expression levels were assessed by Western blotting.

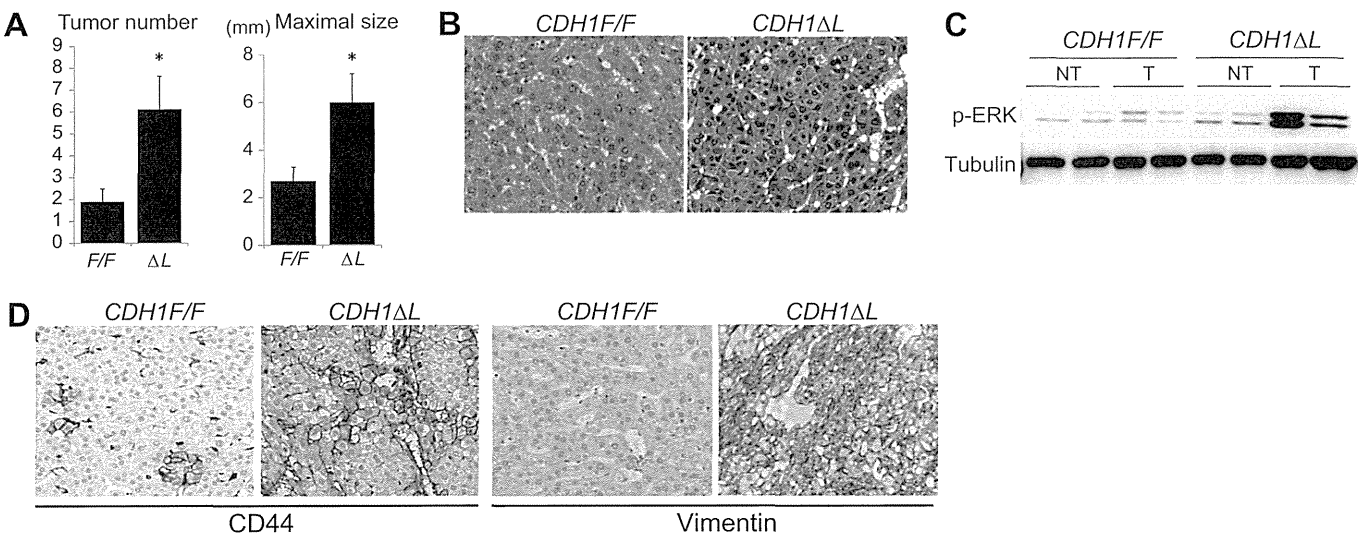

Fig. S7. Loss of E-cadherin promotes DEN-induced HCC. (*A*) *CDH1*$^{F/F}$ (n = 11) and *CDH1*$^{\Delta L}$ (n = 16) mice were injected with 25 mg/kg DEN on postnatal day 14. After 8 mo, tumor number and tumor size were determined. Data are expressed as the means ± SEM, $^{*}P < 0.05$, compared with *CDH1*$^{F/F}$ mice. (*B*) Representative H&E staining of tumors from DEN-treated *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice (×400). (C) Western blotting of ERK phosphorylation in NT and T tissues from *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice. (*D*) Expression of CD44 and vimentin in tumor tissues from *CDH1*$^{F/F}$ and *CDH1*$^{\Delta L}$ mice was assessed by IHC (×400).

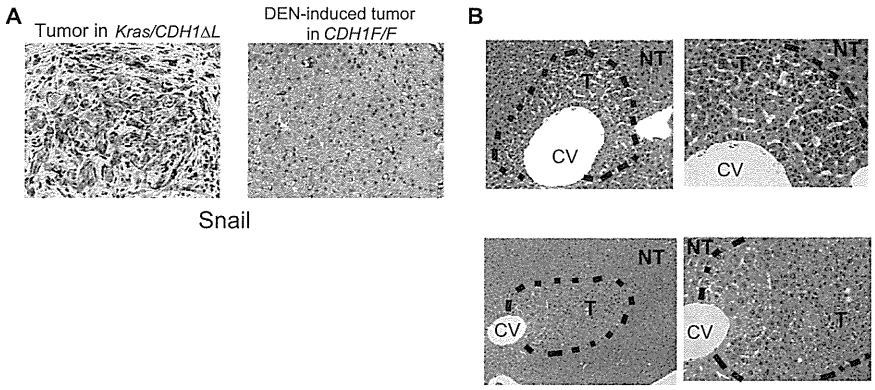

Fig. S8. Characteristics of *Kras/CDH1*$^{\Delta L}$ tumors. (*A*) IHC analysis of Snail in *Kras/CDH1*$^{\Delta L}$ tumors (*Left*, ×400). Tumor cells undergoing epithelial–mesenchymal transition (EMT) strongly expressed Snail. A DEN-induced tumor in *CDH1*$^{F/F}$ mice was used as a negative control (*Right*, ×400). (*B*) Some tumors in *Kras/CDH1*$^{\Delta L}$ mice arise from zone 3. H&E staining images of two representative tumors arising from zone 3 in *Kras/CDH1*$^{\Delta L}$ mice (*Left*, ×200; *Right*, ×400). These tumors are adjacent to the central vein (CV) , which is distant from the portal area.

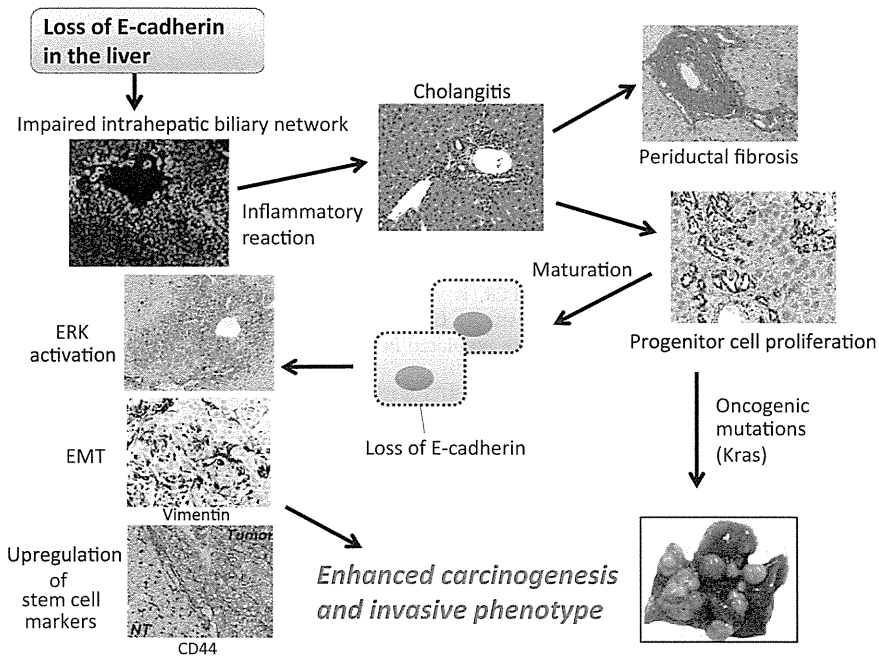

Fig. S9. Proposed mechanism by which loss of E-cadherin induces sclerosing cholangitis and promotes carcinogenesis. Loss of E-cadherin in the liver, especially in BECs, causes impairment of the intrahepatic biliary network and subsequent inflammatory reactions that lead to progenitor cell proliferation and periductal fibrosis. In some cases, these progenitor cells may develop directly into tumors through oncogenic mutations. In mature hepatocytes, loss of E-cadherin leads to ERK activation, EMT induction, and up-regulation of stem cell markers, which eventually results in enhanced carcinogenesis and an invasive phenotype.

Virology 450-451 (2014) 13–23

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

In vitro selection of the 3′-untranslated regions of the human liver mRNA that bind to the HCV nonstructural protein 5B

Kazuhito Yuhashi [a], Shin Ohnishi [a], Tatsuhiko Kodama [b], Kazuhiko Koike [a], Hiroshi Kanamori [a,*]

[a] Department of Gastroenterology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
[b] Laboratory for Systems Biology and Medicine, Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan

ARTICLE INFO

Available online 18 December 2013

Keywords:
HCV
NS5B
LGALS1 (GAL-1)
mRNA 3′-UTR
RPS4X

ABSTRACT

Hepatitis C virus (HCV) nonstructural protein 5B (NS5B) has RNA-dependent RNA polymerase (RdRp) activity. Because NS5B recognizes various RNA motifs besides the HCV genome, NS5B has the potential of interacting with host RNA molecules. In this study, an RNA pool enriched with the 3′-UTR sequences was generated and mRNA molecules with high affinity binding to NS5B were selected by iterative selection. Among the high binding mRNA 3′-UTR segments, we analyzed the housekeeping ribosomal protein S4, X-linked [RPS4X] mRNA 3′-UTR and the 3′-UTR of galectin-1 (GAL-1) mRNA, which is known to be one of the genes upregulated in HCV-infected liver cells and to have a wide spectrum of biological properties. By means of IP-RT-PCR, it was demonstrated that both of the mRNA molecules bind to NS5B in the cytoplasm. Interestingly, GAL-1 and RPS4X mRNA can serve as templates for NS5B RdRp, suggesting these RNA molecules are regulated *in vivo* by NS5B.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Hepatitis C virus (HCV) is a member of the *Flaviviridae* family and contains a positive-sense, single-stranded RNA genome of approximately 9.6 kb (Choo et al., 1989; Kato et al., 1990; Takamizawa et al., 1991), which encodes 4 structural and 6 nonstructural viral proteins (Eckart et al., 1993; Grakoui et al., 1993; Lin et al., 1994; De Francesco 1999). The nonstructural protein 5B (NS5B), which has RNA-dependent RNA polymerase (RdRp) activity (Behrens et al., 1996), has been shown to play a pivotal role in the replication process of the virus. Purified NS5B transcribes the full-length HCV RNA *in vitro* via a copy-back mechanism (Lohmann et al., 1997) and/or *de novo* replication (Luo et al., 2000). Although the crystal structure of NS5B has been determined (Lesburg et al., 1999; Bressanelli et al., 1999; Ago et al., 1999) and the host cellular factors required for HCV replication have been identified (Poenisch and Bartenschlager, 2010), the mechanism of HCV replication is still not completely clear.

The 3′-non-coding region of HCV is composed of variable sequences of approximately 40 nucleotides, a poly(U/UC) tract, and a highly conserved 98-nucleotide 3′-terminal segment termed 3′X (Kolykhalov et al., 1996; Tanaka et al., 1995; Yamada et al., 1996). Earlier studies showed that NS5B binds to both the 3′X RNA and the poly(U/UC) tract (Cheng et al., 1999; Oh et al., 2000). More recent studies have indicated that NS5B binds also to cognate coding sequences (Kim et al., 2002; Lee et al., 2004). Our previous study revealed that one of the stem-loop structures in the NS5B coding region, 5BSL3.2, which is indispensable for viral replication, binds tightly to NS5B (Kanamori et al., 2009). The complimentary strand of 5BSL3.2 RNA, which forms a similar secondary structure to the positive strand, was also shown to bind to NS5B, suggesting the mechanistic importance of the RNA-protein interaction for synthesis of the plus-strand (Kanamori et al., 2010). The 5BSL3.2 RNA sequence interacts not only with the NS5B protein but also with its *cis*-acting elements on the RNA genome, resulting in the formation of a replication initiation complex in association with cellular proteins. The 5BSL3.2 stem-loop structure and the distant stem-loop structure (SL2) in the 3′X region have been shown to form a pseudoknot (Friebe et al., 2005). More recent study has shown that 5BSL3.2 interacts with domain IIId of the internal ribosome entry site located in the 5′-UTR (Romero-Lopez and Berzal-Herranz, 2009).

Beside its own genomic RNA, NS5B has been shown to be able to utilize diverse RNA molecules such as rat dimerization cofactor of hepatocyte nuclear factor-1α (Behrens et al., 1996), various homopolymers (Lohmann et al., 1997) and certain synthetic RNA molecules (Kao et al., 2000) as its RdRp templates. On the other hand, certain template RNA sequence requirement for NS5B to synthesize the anti-strand RNA does appear to exist. Strict

* Corresponding author. Tel./fax: +81 3 5800 8901.
E-mail address: hkanamori-tky@umin.net (H. Kanamori).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.11.036

sequence-specific RdRp activities of NS5B on RNA molecules have been demonstrated in experiments using artificial RNA stem-loop structures (Kanamori et al., 2009; Kao et al., 2000; Biroccio et al., 2002; Vo et al., 2003).

Thus, NS5B can bind to or replicate several different RNA motifs depending on the sequence and structure of the RNA. In virus-infected cells, certain viral proteins such as the SOX protein Kaposi's sarcoma-associated herpes virus and the vaccinia D10 protein regulate cellular RNA expression through direct interactions (Lee and Glaunsinger, 2009; Parrish et al., 2007). The wide spectrum of NS5B RNA recognition suggests the possible interaction of NS5B with cellular mRNA molecules and regulation of these mRNA species by NS5B in the infected cells. In fact, the HCV products appear to induce an increase or decrease in the expression levels of cellular genes associated with cellular defense mechanisms, cellular metabolism or intracellular transport (Blackham et al., 2010). Among these genes, GlcT-1 (ceramide glucosyltransferase) expression is known to be induced by NS5B (Guo et al., 2012).

The mRNA 3′-UTRs have been intensively investigated in studies of the RNA binding proteins, in part because the mRNA 3′-UTRs are more accessible to cellular factors than are the coding regions of mRNA, which usually form polyribosomes in association with ribosomes (Warner et al., 1963; McCarthy and Kollmus, 1995; Levine et al., 1993; Antic et al., 1999; Imai et al., 2001).

In the present study, we constructed an RNA library that mainly contained the 3′-UTR portions of human liver cell line mRNA, selected RNA species that were found to bind tightly to the NS5B

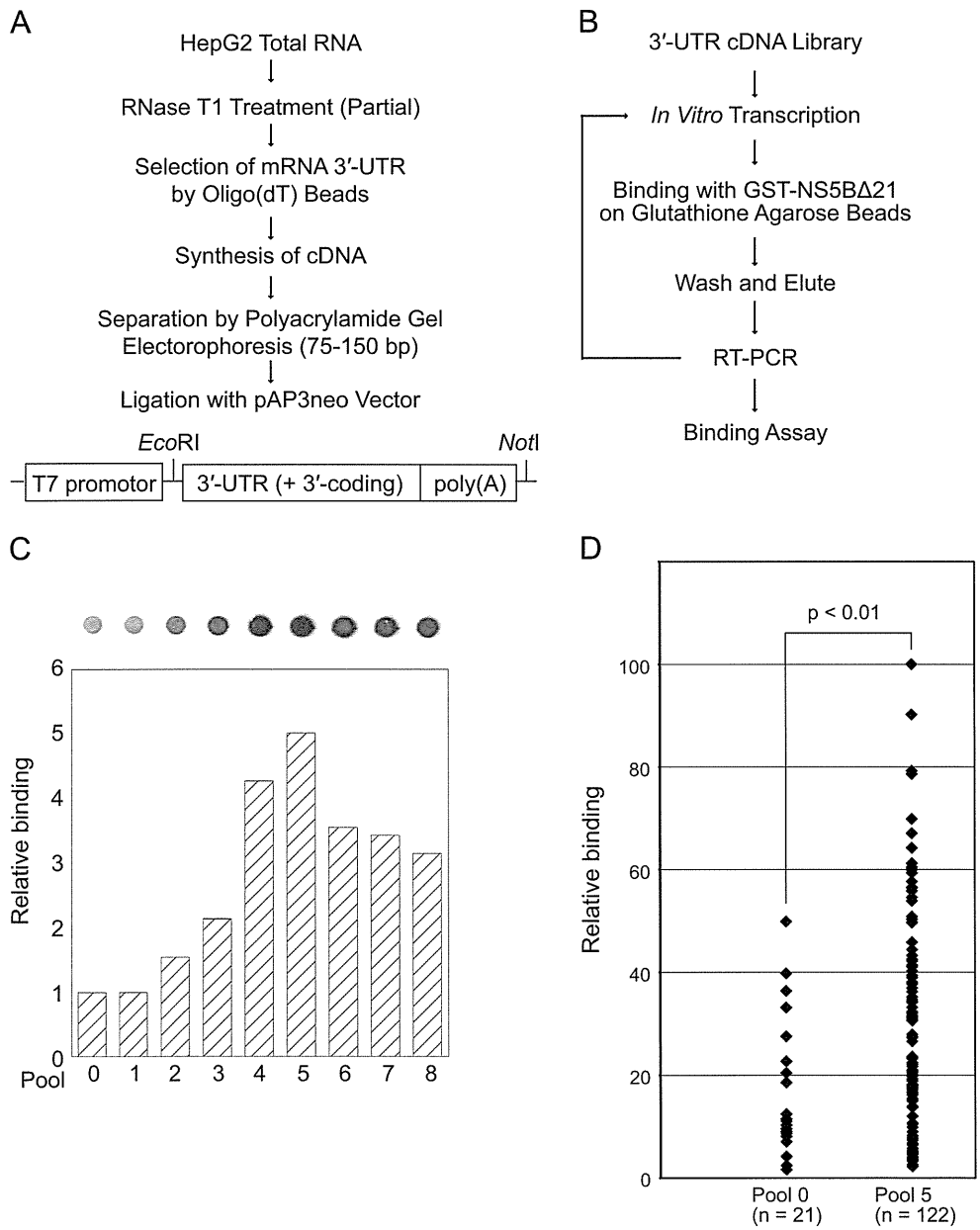

Fig. 1. Outline of the *in vitro* selection for high affinity binding hepatic mRNA 3′-UTR sequences to the HCV NS5B protein. (A) Construction of a hepatic cDNA library enriched with 3′-UTR sequences. The bottom portion indicates the construction of the template plasmids for the mRNA 3′-UTR library. (B) Scheme for *in vitro* selection of the mRNA 3′-UTRs that exhibit high-affinity binding to NS5B. (C) Relative binding to NS5B was determined by filter binding analysis and is represented by a bar graph. The dot blot analysis of the RNA pools from the different selection cycles is shown at the top. The relative binding of pool 0 was set at 1. (D) The relative binding strength of the individual clones from pool 0 ($n=21$) and pool 5 ($n=122$) was determined by filter binding dot blot analysis. The binding strength of each clone was plotted relative to that of the highest binder, RPS5 3′-end mRNA, which was set at 100. The significance of the difference in binding ability to NS5B between pools 0 and 5 RNA clones was analyzed by Welch's *t*-test.

protein of HCV by employing iterative selection cycles, and identified mRNA 3′-UTRs that can be templates of the NS5B polymerase.

Results

Selection of mRNA 3′-UTR sequences that bind tightly to the HCV NS5B protein

We constructed a cDNA library that mostly consisted of 3′-UTR sequences plus poly(A) tails (Fig. 1A). The 3′-UTR mRNA portion varies in length, ranging from less than 50 to over 5,000 nucleotides, with many of the mRNA 3′-UTRs around a 100-nucleotide length (Grillo et al., 2010). To construct a cDNA library enriched with the 3′-UTR sequences, we added a step in which partial digestion of the total RNA from HepG2 cells with RNase T1 was included. After the synthesis of cDNA from the partially digested liver cell poly(A) RNA, the cDNA with a length between 75 and 150 nucleotides was constructed under the T7 promoter to generate the templates for the initial RNA library for the selection. We estimated that the library contained approximately 10^5 independent clones. By sequence analysis of the library (40 clones), we found that the cDNA clones contained 3′-UTRs of 15–105 nucleotides in length (on average 56.2 nucleotides) and the poly(A) tails (on average 18 nucleotides). The initial RNA pool (pool 0) was synthesized by *in vitro* transcription using the cDNA library as a template.

By applying a selection cycle using GST-NS5B (genotype 1b, strain BK) as the target molecule (Fig. 1B), we obtained RNA clones with stronger binding to NS5B. In the selection procedure, the bound poly(A)+ RNA molecules to NS5B were recovered, amplified by RT-PCR, *in vitro* transcribed and subjected to iterative selection cycles. The binding capacities of each RNA pool was

Table 1
Genes for which the 3′-RNA sequences bind tightly to HCV NS5B.

Symbol	Accession	Gene name	Expression level[a]	Relative binding[b]	
				clone	3′-UTR
RPS5	NM_001009.3	Ribosomal protein S5	2460.5	100	68
ATP5I	NM_007100.2	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E, nuclear gene encoding mitochondrial protein	470.6	90	ND
C5	NM_001735.2	Complement component 5	189.5	79	ND
GNB2L1	NM_006098.4	Guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	2869.2	70	96
NFS1	NM_021100.3	NFS1 nitrogen fixation 1 homolog (S. cerevisiae), nuclear gene encoding mitochondrial protein	155.4	67	ND
RPS15A	NM_001019.4	Ribosomal protein S15a, transcript variant 2	4826.1	64	33
RPS4X	NM_001007.3	Ribosomal protein S4, X-linked	3404.1	61	85
RPL41	NM_001035267.1	Ribosomal protein L41, transcript variant 2	5374.1	60	93
LGALS1 (GAL-1)	NM_002305.3	Lectin, galactoside-binding, soluble, 1 (galectin-1)	80.1	59	100

[a] The relative expression level of each gene in HepG2 human hepatoblastoma cells was determined using CHIP (HG-U133A) from the Reference database for gene expression analysis: RefExA (http://157.82.78.238/refexa/main_search.jsp).
[b] The relative binding strength of each pool 5 RNA clone to GST-NS5B was determined by dot blot experiments. The binding strength of the RPS5 clone was set equal to 100. The relative binding strength of the 3′-UTR RNA sequence (47 nucleotides) of each gene to GST-NS5B was determined by the RNA gel mobility shift analysis shown in Fig. 2A. The binding strength of the GAL-1 3′-UTR was set equal to 100.

Table 2
Frequency of each selected gene in pool 0 and pool 5.

Gene	Nucleotide length[a]		Frequency		Functional category
	Coding	3′-UTR	Pool 0	Pool 5	
RPS5	30 (615)	60	1/67	1/188	Structural constituent of ribosome
ATP5I	10 (210)	61	0/67	1/188	Contributes to ATPase activity, hydrogen ion transmembrane transporter activity, transmembrane transporter activity
C5	0 (5031)	72 (401)	0/67	1/188	C5a anaphylatoxin chemotactic receptor binding, chemokine activity, endopeptidase inhibitor activity, receptor binding
GNB2L1	39 (954)	21	0/67	2/188	GTPase activity, protein kinase C binding, receptor activity, kinase activity
NFS1	0 (1374)	93 (657)	0/67	1/188	Cysteine desulfurase activity, protein binding, protein homodimerization activity, pyridoxal phosphate binding
RPS15A	51 (393)	41	0/67	5/188	structural constituent of the ribosome
RPS4X	4 (792)	64	0/67	4/188	RNA binding, rRNA binding, structural constituent of the ribosome
RPL41	0 (78)	79 (308)	1/67	3/188	Structural constituent of the ribosome
LGALS1 (GAL-1)	44 (408)	51	0/67	1/188	Galactose binding, promotes apoptosis of activated T cells, regulation of CD4+CD25+ T cells, tumor-immune escape and correlation with tumor aggressiveness
Total			2/67 (0.030)	19/188 (0.101)	

[a] Nucleotide length of coding and 3′-UTR lesions of each selected clone is indicated. In the parenthesis, the nucleotide length of the full length coding or 3′-UTR of each cDNA is indicated.

examined by dot blot analysis (Fig. 1C), which indicated that the binding capacity of the RNA pools was enriched in the earlier cycles of the selection. The RNA in pool 5 appeared to show the strongest binding to NS5B.

Individual RNA clones from pool 5 exhibited significantly stronger binding to NS5B compared to those from pool 0 ($P<0.01$, Fig. 1D). The 9 highest binding mRNA 3′-UTR clones in pool 5 are listed in Table 1. The C5, NFS1 and RPL41 clones contained the 3′-end of the 3′-UTR sequences (Table 2), while the other 6 clones contained the full length 3′-UTR sequences plus the 3′-end of the coding sequences. A consensus RNA motif was not identified among these clones even using the computer program Nucleotide BLAST.

We analyzed the binding capacity of the 3′-UTR sequences from the pool 0 and 5 clones by gel mobility shift assay (Fig. 2A). Because these clones contained various lengths of the cDNA fragments under the T7 promoter, in this experiment, the 47-nucleotide RNA sequence just upstream of the 3′-end of each clone was synthesized and used in the gel mobility shift assay. The RNA-protein complexes appeared to form multiple bands, representing the oligomeric character of the NS5B protein (Lee et al., 2004). The pool 0 RNA 3′-UTRs exhibited relatively poor binding to NS5B. While the pool 5 clone RPS15A 3′-UTR exhibited poor binding, 5 out of 6 of the 3′-UTR sequences from pool 5 exhibited tighter binding than those from pool 0.

The 3′-UTRs of GAL-1 and RPS4X mRNA bind tightly to NS5B

We further analyzed two of the mRNA clones. GAL-1: lectin, galactoside-binding, soluble, 1 (LGALS1) mRNA encodes the galectin-1 protein, the expression level of which in HepG2 cells is known to be relatively low (Table 1). RPS4X: ribosomal protein S4, X-linked mRNA, is expressed at high level in a wide range of organs (Table 1, http://157.82.78.238/refexa/main_search.jsp). These two genes have relatively short 3′-UTRs (51 and 68 nucleotides, respectively), which we would make them suitable for further analysis. A gel mobility shift assay revealed that both the 3′-UTRs of GAL-1 mRNA and RPS4X mRNA formed a complex, not only with GST-NS5B, but also with NS5B and not GST (Fig. 2B and C), indicating specific binding of the RNA elements to NS5B. The secondary structure analysis by the mfold computer program (Zuker, 2003) indicated the presence of 2 stem-loop structures in the 3′-UTR of GAL-1 mRNA (Fig. 3A, GAL-1 1–51), while the 3′-UTR of RPS4X mRNA did not have any evident secondary structure and appeared to be single-stranded (Fig. 3C).

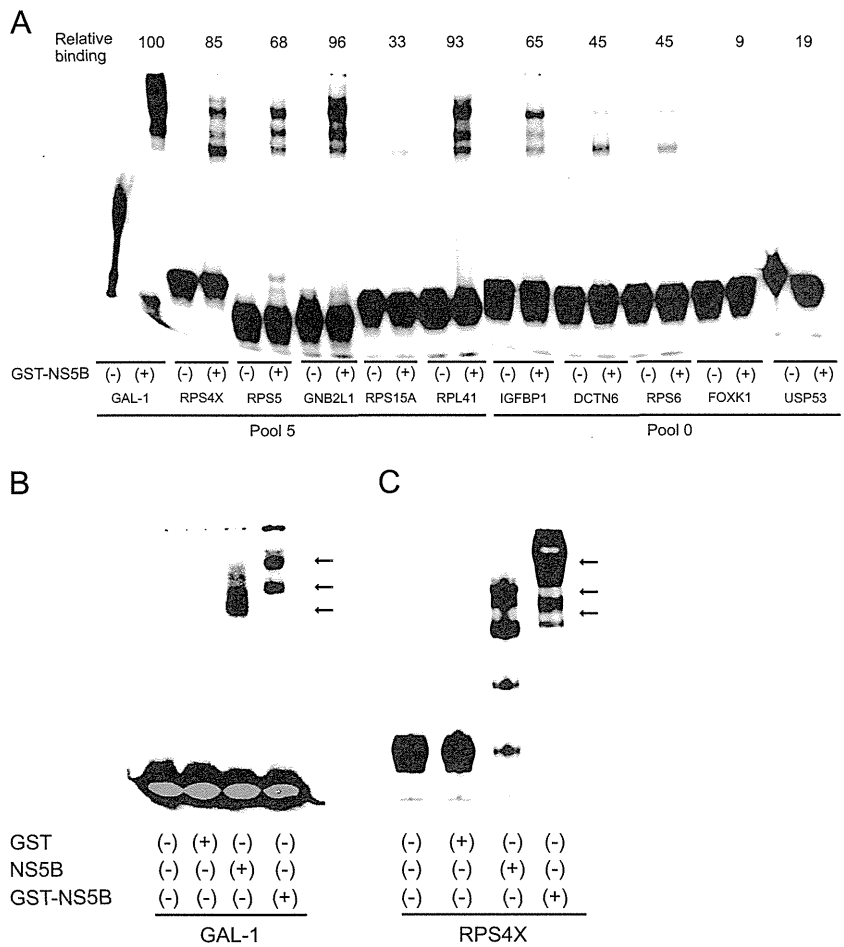

Fig. 2. RNA gel mobility shift assay using 3′-UTR mRNA sequences of the pool 0 and pool 5 clones. (A) An RNA segment (47-nucleotide length) just upstream of the 3′-end of each clone from pool 0 and pool 5 was synthesized, 5′-end labeled with 5′-[^{32}P] and incubated with (+) or without (−) GST-NS5B, and mixtures were run on a 5% polyacrylamide gel. The relative binding strength of each clone was determined by quantification by BAS2000 and is shown on the top portion of the gels. The binding strength of GAL-1 was set equal to 100. IGFBP1: insulin-like growth factor binding protein 1, DCTN6: dynactin 6, RPS6: ribosomal protein S6, FOXK1: forkhead box K1, USP53: ubiquitin specific peptidase 53. The specific interaction of the 3′-UTRs of GAL-1 and RPS4X mRNA with the NS5B protein is shown in panel B and C, respectively. GST, GST-truncated NS5BΔ21 (NS5B) or the GST-NS5BΔ21 (GST-NS5B) protein was incubated with either the GAL-1 (GAL-1 1–51, panel B) or RPS4X (RPS4X 1–68, panel C) RNA 3′-UTR as a probe and analyzed by RNA gel mobility shift assay. The arrows indicate the positions of the RNA-protein complexes.

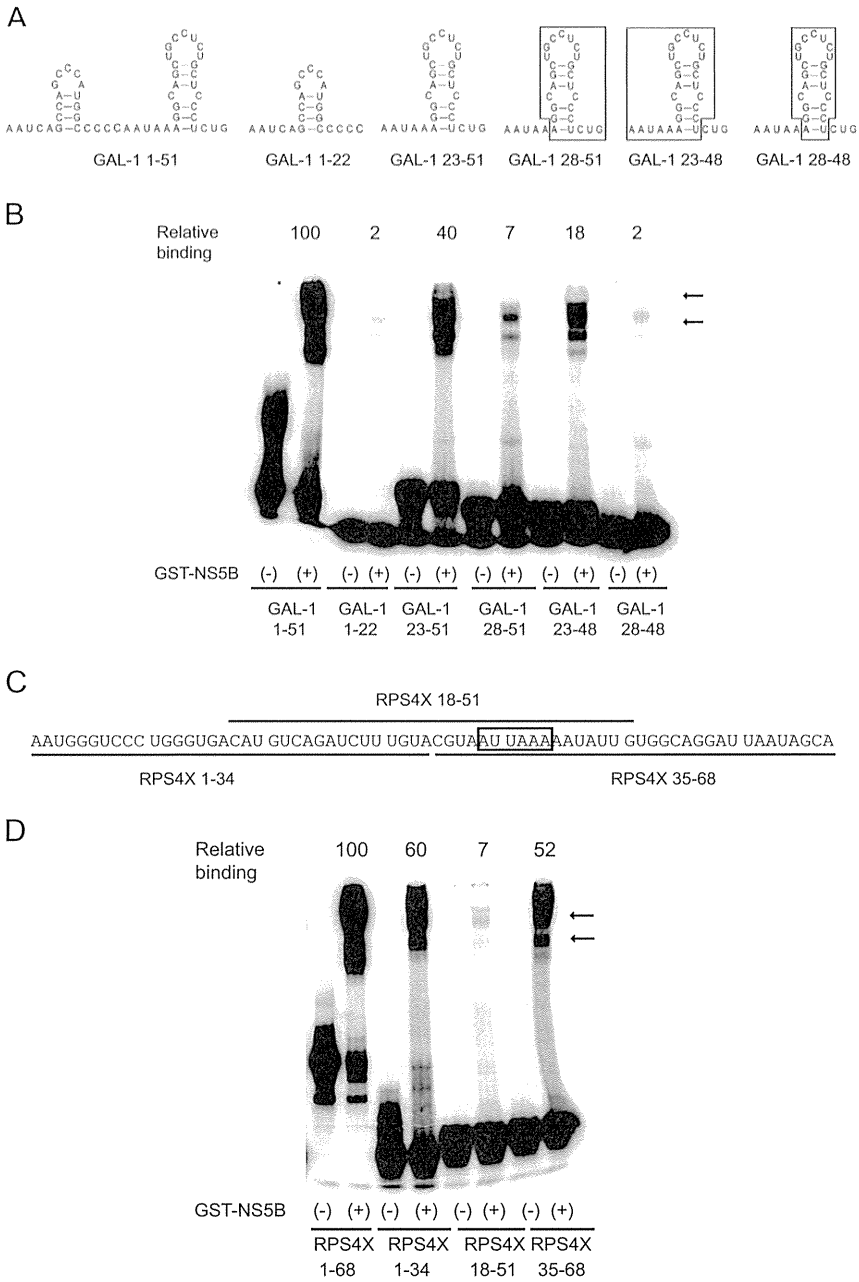

Fig. 3. Analysis of the deletion mutants of GAL-1 and RPS4X mRNA 3′-UTRs by RNA gel mobility shift assay using the NS5B protein. (A) Predicted secondary structures of GAL-1 1–51 and the deleted RNA clones (1–22, 23–51, 28–51, 23–48, and 28–48) by mfold computer analysis (Zuker, 2003) are shown. For GAL-1 28–51, 23-48 and 28–48, each secondary structure is shown in a box. (B) An RNA gel mobility shift assay using GAL-1 3′-UTR (1–51) and the deleted clones (1–22, 23-51, 28–51, 23–48 and 28–48). Chemically synthesized RNA was labeled with 5′-[^{32}P], incubated with (+) or without (−) GST-NS5B, and mixtures were run on a 5% polyacrylamide gel. (C) Nucleotide sequence of RPS4X 3′-UTR (1–68) RNA is shown. The 34 nucleotide RNA fragments (1–34, 18–51, 35–68) tested in panel D are indicated. The putative polyadenylation signal is boxed. (D) An RNA gel mobility shift assay using RPS4X 3′-UTR (1–68) and its deleted sequences (1–34, 18-51 and 35–68) in the presence (+) and absence (−) of the NS5B protein. The relative binding strength of each RNA molecule is shown at the top of the gels. The binding strength of GAL-1 1–51 (panel B) or RPS4X 1–68 (panel D) was set equal to 100. The arrows indicate the positions of the RNA-NS5B complexes.

Both the stem-loop structures and the single-strand regions of the GAL-1 3′-UTR are required for binding to NS5B

We undertook mapping of the NS5B binding sites on the 3′-UTRs of GAL-1 and RPS4X mRNA by testing the binding strength of the various portions of the 3′-UTR RNA segments by RNA gel mobility shift analysis, as shown in Fig. 3. Full length GAL-1 3′-UTR (GAL-1 1–51) RNA exhibited strong binding to NS5B (Fig. 3B). The 3′-half of the GAL-1 3′-UTR (GAL-1 23–51) exhibited weaker but still substantial binding, while the 5′-half (GAL-1 1–22) did not form a complex with NS5B. RPS4X 1–68 exhibited strong binding to NS5B. Both the 5′- (RPS4X 1–34) and the 3′-half (35–68) displayed weaker but still substantial levels of binding strength to NS5B (Fig. 3D). Interestingly, the middle portion of the RPS4X 3′-UTR (18–51) did not exhibit any binding to GST-NS5B. Because the nucleotide length of the 3 RPS4X RNA fragments tested were the same (34 nucleotides), the binding strength is not likely to be determined by the nucleotide length. The result instead indicates that either the 5′- or 3′- end of the 3′-UTR of RPS4X mRNA is indispensable for tight binding.

From the results of the GAL-1 mRNA binding, we thought that the stem-loop structure located on the 3′-half of the GAL-1 3′-UTR

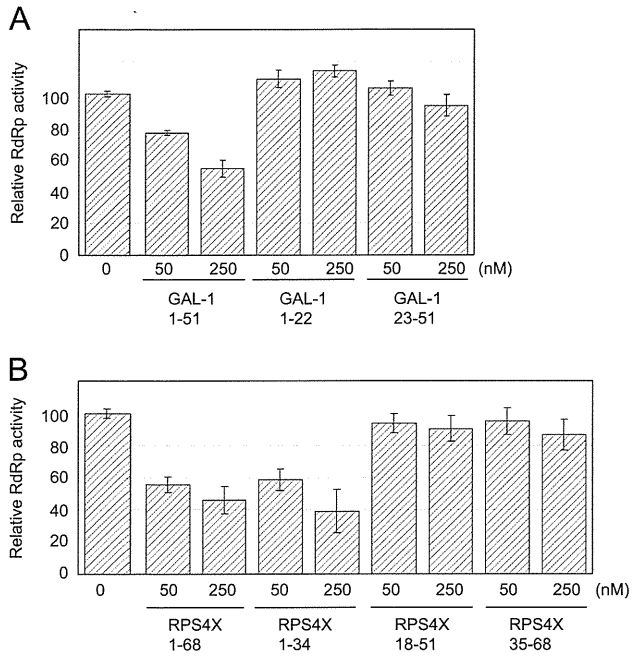

Fig. 4. Inhibition of RdRp activities by the mRNA 3′-UTRs. (A) GAL-1 mRNA or one of its deleted sequences (GAL-1 1–22 and 23–51, upper panel) was added in the poly(C)-oligo(G) RdRp reaction mixture (0, 50, and 250 nM) and the RdRp activity was measured. (B) The 3′-UTR of RPS4X mRNA or one of its deleted sequences (RPS4X 1–34, 18–51 and 35–68) was used as a competitor (0, 50, and 250 nM). The experiment was performed in triplicate. The standard deviations are shown by the error bars.

might be the key element in the tight binding of RNA to NS5B. The stem-loop structure is flanked by the 6-nucleotide polyadenylation signal (AAUAAA) on the 5′-end and the 3 nucleotide stretch (CUG) just upstream of the poly(A) tail on the 3′-end. To further analyze the binding of the GAL-1 3′-UTR RNA sequence to NS5B, we tested the deleted RNA oligonucleotides (GAL-1 28–51, 23–48 and 28–48, Fig. 3A) for the capacity to bind NS5B (Fig. 3B, right lanes). Interestingly, deletion of either of the flanking single-stranded nucleotides (GAL-1 28–51, 23–48) greatly reduced the capacity of the RNA to bind to NS5B. The stem-loop structure by itself (GAL-1 28-48) did not bind to NS5B.

mRNA 3′-UTRs with strong binding to NS5B inhibit the NS5B RdRp activity measured by poly(C)-oligo(G) system

The *in vitro* RdRp assay is a versatile assay that is used to evaluate the activity of the inhibitory factors against the NS5B polymerase (Kanamori et al., 2010). We used a primer-dependent RdRp system to evaluate the inhibitory effects of RNA segments from the 3′-UTRs of GAL-1 and RPS4X mRNA on NS5B RdRp. The full length 3′-UTRs of GAL-1 (GAL-1 1–51), RPS4X (RPS4X 1–68), or one of the dissected RNA fragments was added (50 or 250 nM) to the RdRp activity assay (Fig. 4). The full length 3′-UTRs of GAL-1 and RPS4X inhibited the RdRp activities efficiently. While RPS4X 1–34 also reduced the RdRp activity to 40% (Fig. 4B), none of the other dissected 3′-UTR RNA fragments (GAL-1 1–22, 23–51, RPS4X 18–51 or 35–68) inhibited the activity effectively.

The 3′–UTR sequences that bind to NS5B can serve as templates of NS5B RdRp

We tested whether the mRNA 3′-UTR molecules that exhibited potent binding to NS5B can serve as substrates of NS5B RdRp by adding RNA fragments to the NS5B RdRp reaction mixture (Fig. 5). In addition to pool 5 (GAL-1 and RPS4X) RNA molecules, pool 0 FOXK1 3′-UTR was very efficiently elongated when the poly (A) tail was present at the 3′-end (Fig. 5A).

Next, we tested whether the 3′-UTR sequence without the poly (A) tails can serve as a template for NS5B RdRp (Fig. 5B). The full length 3′-UTRs of GAL-1 (GAL-1 1–51) and RPS4X (RPS4X 1–68) appeared to be elongated by NS5B efficiently, while the 3′-end of FOXK1 3′-UTR, which is one of the pool 0 clones that displayed poor binding to NS5B (Fig. 2), was not elongated. In correlation with the binding ability (Fig. 3B), the 5′-portion of the GAL-1 3′-UTR (1–22) was elongated less efficiently compared with its 3′-portion (23–51).

The binding of GAL-1 and RPS4X mRNA to NS5B in vivo

To investigate whether the GAL-1 and RPS4X mRNA sequences bind to NS5B *in vivo*, we performed immunoprecipitation (IP)–RT-PCR assay (Fig. 6). RCYM1 cells are derived from Huh7 cells and harbor a genome-length dicistronic HCV genotype 1b full-length RNA. In this experiment, RCYM1 cells were cultured in a monolayer, where the viral proteins are efficiently produced while the viral particles are less efficiently secreted compared with those in three-dimensional culture (Murakami et al., 2006). A cytoplasmic extract of the monolayer-cultured RCYM1 cells was immunoprecipitated with a NS5B antibody in order to monitor the binding of RNA molecules to NS5B in the HCV replication complex by applying RT-PCR to the co-immunoprecipitated RNA molecules using specific primer pairs (Fig. 6A). Huh7 parental cells appeared to contain substantial amounts of the tested RNA molecules (Fig. 6A, GAL-1, RPS4X and FOXK1, Total RNA, Huh7). The RT-PCR product from GAL-1 as well as from RPS4X RNA was detected in the IP eluate of RCYM1 cells, while it was not detected in the parental Huh7 cell IP eluate, indicating the *in vivo* interaction of these RNA molecules with NS5B in the replication complex. FOXK1 RNA, which does not bind to NS5B, was not detected in the IP eluate of either RCYM1 cells or Huh7 cells, in spite of the fact that the total Huh7 cell RNA appears to contain a substantial amount of the FOXK1 RNA (Fig. 6A, FOXK1, Total RNA, Huh7). The results indicate that NS5B as well as other components in the HCV replication complex interact specifically with the GAL-1 and RPS4X RNA molecules.

To monitor the NS5B-RNA interaction *in vivo* more specifically, we tested COS cells transiently transfected with the FLAG-fused NS5B gene (Fig. 6B). The FLAG-fused NS5B protein from the cytoplasmic lysates of COS cells transfected with an NS5B expressing plasmid (pFLAG-5B) was immunoprecipitated with a FLAG antibody. RT-PCR was performed to monitor the specific binding of RNA molecules to NS5B. Non-transfected COS cells appeared to contain substantial amounts of the tested RNA molecules (Fig. 6B, GAL-1, RPS4X and FOXK1, Total RNA, Transfection (−)). The RT-PCR products from GAL-1 RNA and RPS4X RNA were detected in the IP eluate of COS cells expressing the NS5B protein (pFLAG-5B), while they were not detected in the mock–transfected COS cell IP eluate (pFLAG-CMV-4). In contrast, FOXK1 RNA, which does not bind to NS5B, was not detected in the IP eluate from the NB5B expressing COS cells (pFLAG-5B, FOXK1). These results indicate that the GAL-1 and RPS4X mRNA molecules directly interact with NS5B *in vivo*.

Discussion

We constructed a cDNA library that contained mostly 3′-UTR sequences from a human hepatoblasma cell line, HepG2. By using iterative selection cycles, we obtained 3′-UTR RNA segments that bind tightly to the HCV NS5B protein. In an earlier study by Gao et al. (1994), a similar cDNA library was constructed using a simpler method and 3′-UTR sequences were identified that bind

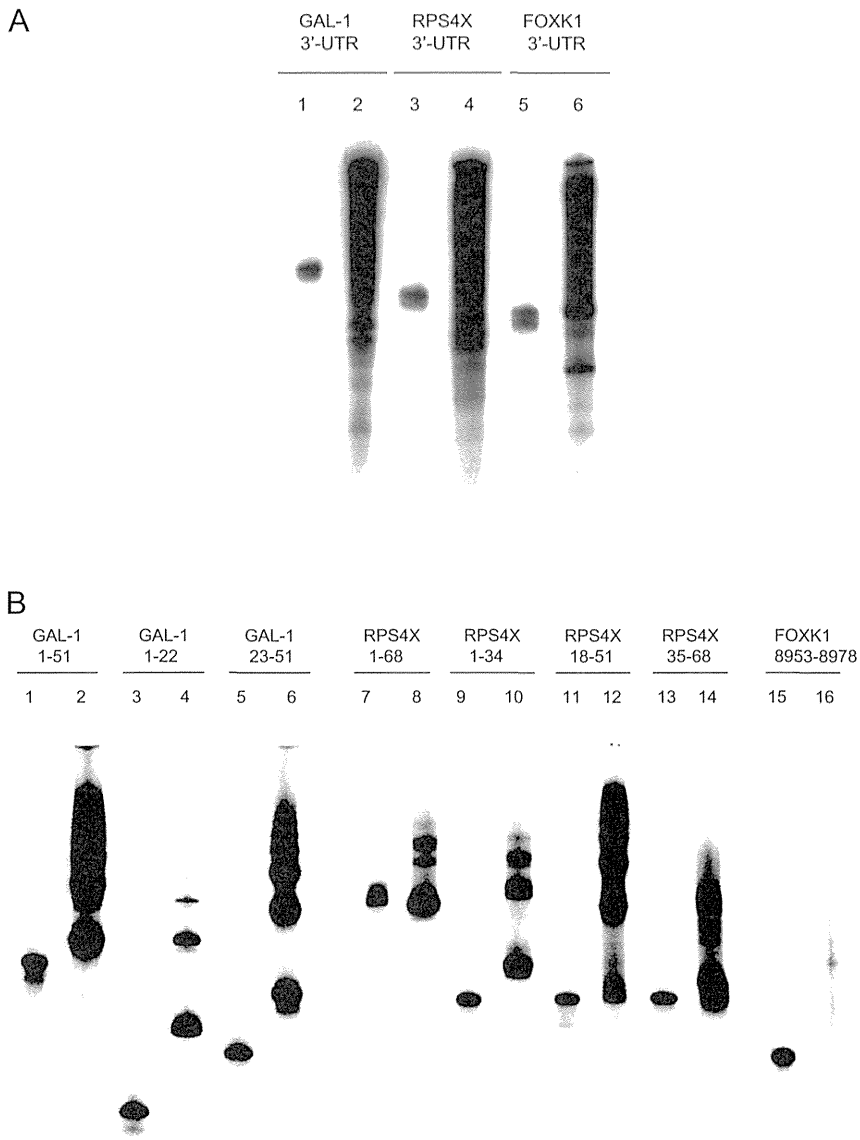

Fig. 5. RNA synthesis by NS5B RdRp using RNA 3′-UTR sequences as the templates. (A) The 3′-portion of GAL-1 (the 44-nucleotide coding, 51-nucleotide 3′-UTR and 26-nucleotide poly(A) tail), RPS4X (the 4-nucleotide coding, 64 nucleotide 3′-UTR and 24-nucleotide poly(A) tail) or FOXK1 (the 59-nucleotide 3′-UTR and 18-nucleotide poly (A) tail) mRNA segment was incubated with recombinant NS5B, cold NTPs and [α-^{32}P]UTP, and the RdRp products were analyzed by electrophoresis on a 12% polyacrylamide/ 7 M urea gel. For reference, the RNA templates were labeled at the 5′-end with [γ-^{32}P] by using T4 polynucleotide kinase and electrophoresed side-by-side with each RdRp product (lanes 1, 3 and 5). B. The 3′-UTR sequences without any poly(A) tails were tested similarly. The 3′-UTR of GAL-1 (GAL-1 1–51), its deleted form (GAL-1 1–22, 23–51), the 3′-UTR of RPS4X (RPS4X 1–68), its deleted form (RPS4X 1–34, 18–51, 35–68) or the FOXK1 3′-UTR segment (3′-end 26 nucleotides) was tested as a template for NS5B RdRp. For reference, the RNA templates were labeled at the 5′-end with [γ-^{32}P] using T4 polynucleotide kinase and electrophoresed side-by-side with each RdRp product (lanes 1, 3, 5, 7, 9, 11, 13 and 15).

to Hel-N1/Hel-N2 proteins. In the present study, by performing partial RNase T1 treatment of HepG2 RNA and size fractionation of cDNA fragments, we successfully removed approximately 99% of the coding sequences from the cDNA library. According to the UTR data base, approximately 50% of the 3′-UTRs of mRNA are less than 500 nucleotides (Grillo et al., 2010). Therefore, we decided to focus our attention on the shorter 3′-UTR sequences and recovered cDNA fragments 75–150 nucleotides in length, comprising approximately 9% of mRNA 3′-UTR sequences. As a result, 67.5% (27/40) of the clones in the library contained the 3′-end of the 3′-UTR (an average of 60.4 nucleotides) plus poly(A) tails and did not contain any coding regions. The rest of the pool 0 clones (13/40) contained the full length 3′-UTRs (an average of 47.5 nucleotides) and very short coding sequences (an average of 36.2 nucleotides) plus poly(A) tails.

The binding capacity of the RNA pool to NS5B increased as the selection cycles proceeded, reaching a peak at pool 5. Among the highest 9 binders in pool 5, 4 were housekeeping genes (Table 1). The frequency of these 9 high binding clones as a whole in the pool increased after the selection cycles (pool 5, 19/188, Table 2), while there were fewer in pool 0 (2/67), indicating the effectiveness of the selection process ($P < 0.05$).

Stem-loop structures have been considered to be important targets of RNA binding proteins (Dingwall et al., 1989; Koeller et al., 1989;

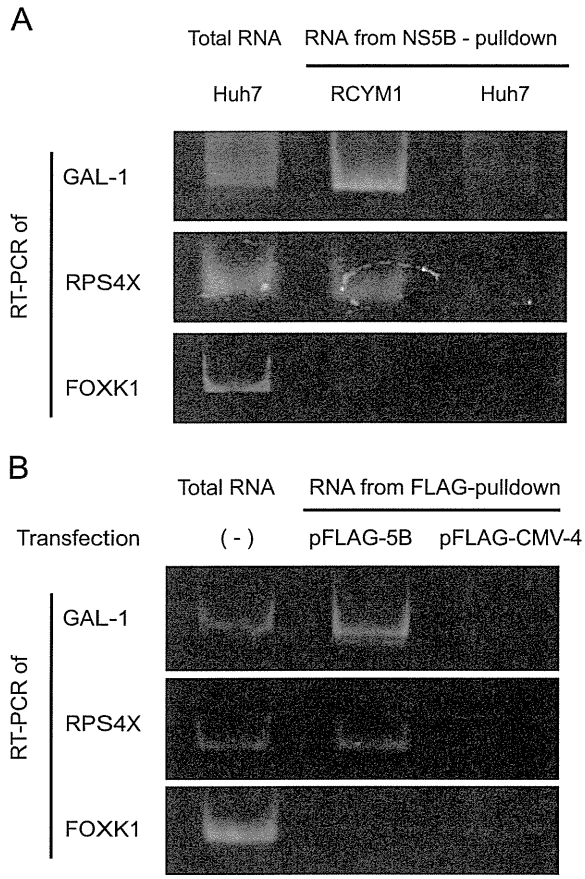

Fig. 6. The specific interaction of the NS5B protein with GAL-1 and RPS4X mRNA 3′-UTRs *in vivo*. (A) Immunoprecipitation using an anti-NS5B antibody was performed using cell extracts of RCYM1 cells or Huh7 cells. Following the RNA extraction of the immunoprecipitants, RT-PCR was performed using a primer pair specific for each mRNA to monitor GAL-1, RPS4X or FOXK1 mRNA bound to NS5B. As a reference, RT-PCR was carried out using total RNA from Huh7 cells (Total RNA) using the same primer pairs and the PCR products were run on 6% polyacrylamide gels side-by-side with the IP-RT-PCR products. (B) Immunoprecipitation using an anti-FLAG antibody was performed using cell extracts of COS cells transfected with either a FLAG-tag fused NS5B expression vector (pFLAG-5B) or a blank vector (pFLAG-CMV-4). RNA molecules bound to NS5B *in vivo* were amplified by RT-PCR and visualized as described in panel A. As a reference, RT-PCR was carried out using total RNA from nontransfected COS cells (Total, Transfection (−)) using the same primer pairs.

Munishkin et al., 1991; Tan et al., 1996; Nomaguchi et al., 2004; Kim et al., 2007). While the 3′-UTR of RPS4X mRNA is single-stranded and the full length 3′-UTR (68 nucleotides length) appears to be important for tight binding to NS5B, the stem-loop structure at the 3′-end of GAL-1 RNA also appears to be important for NS5B binding. Of note, the flanking 5′- and 3′-single-strand nucleotides were also indispensable for the tight binding to NS5B, having a similar property as the aptamers to NS5B reported in our earlier study (Kanamori et al., 2009).

GAL-1 1–51, as well as RPS4X 1–68 and 1–34 exhibited comparatively high affinity binding to NS5B in gel mobility shift analysis (Fig. 3), and at the same time strong inhibitory effects against RdRp activity (Fig. 4). The binding strength to NS5B was well correlated with the inhibitory effect on NS5B RdRp and also the efficiency as a template of NS5B RdRp in most of the RNA fragments we tested. RPS4X 18–51 exhibited weak binding to NS5B and weak inhibition on the RdRp, and yet was still efficiently elongated by NS5B RdRp. This suggests that tight binding is not necessarily required for the RNA molecules to interact with the NS5B RdRp catalytic site for replication. This result also suggests that there may be a population of mRNA molecules that did not survive the *in vitro* selection in this study but still can be elongated efficiently by NS5B RdRp.

Although limited to the three mRNA clones tested in this study, all of the RNA molecules bearing poly(A) tails on the 3′-ends were very efficiently elongated by NS5B, while the elongation efficiency of the 3′-UTRs without any poly(A) tails depended on their individual sequences (Fig. 5). The *in vivo* situation may be different because the poly(A) tails are likely to form complexes with other cellular factors such as the poly(A) binding protein and the poly (A) polymerase, and the importance of the poly(A) tail in the RdRp reaction should not be underestimated, while the 3′-UTR portions of mRNA appear to be more likely to interact with cellular or viral proteins.

We employed immunoprecipitation (IP)–RT-PCR and demonstrated that the GAL-1 and RPS4X mRNA 3′-UTR sequences actually do interact with NS5B *in vivo*. Notably, these mRNA molecules were shown to bind to NS5B in the replicon cell line RCYM1, where NS5B is one of the components of the HCV replication complex in the ER-derived membrane spherules, which are termed membranous web. The *in vivo* interaction of NS5B and these mRNA molecules suggests a mechanism by which the gene expression levels might be regulated by NS5B in HCV-infected cells. Compared to the abundant, constitutively-expressed housekeeping RPS4X gene product (Eisenberg and Levanon, 2003), the regulation of the gene expression of GAL-1 is biologically of greater importance, as the expression level of the GAL-1 gene is relatively low in the liver (Camby et al., 2006). In the COS cell experiment, where only NS5B was expressed transiently (Fig. 6B), the *in vivo* interaction of these RNA molecules with NS5B was shown to probably occur through a direct interaction with NS5B. The interaction of the NS5B protein with the GAL-1 mRNA 3′-UTR may influence the mRNA turnover rate and increase or decrease its expression level, since the control of the mRNA half life is considered to be regulated mainly by the 3′-UTR portions of the RNA (Ross 1995).

Although percentage of the immunoprecipitated GAL-1 and RPS4X mRNA was estimated to be as little as 0.1% of the total cellular mRNA molecules (Fig. 6), the *in vivo* interaction of NS5B with cellular mRNA may play a role in the host immune response against HCV infection. Generation of double stranded RNA molecules following the elongation of the host mRNA molecules by NS5B RdRp may increase the innate immunity against HCV infection (Ranjith-Kumar et al., 2011).

Recently, GAL-1 has received considerable attention because of its pleiotropic functions, such as promoting the apoptosis of activated T cells (Perillo et al., 1995; Matarrese et al., 2005), cytokine production (IL-5, IL-10, TGF-β), the regulation of CD4+ CD25+ T cells (Santucci et al., 2000; Garin et al., 2007), cell-extracellular matrix and cell-cell interaction, cell migration, angiogenesis, tumor-immune escape and correlation with tumor aggressiveness, especially hepatocellular carcinoma (Liu and Rabinovich, 2005; Rabinovich, 2005; Salatino et al., 2008; Spano et al., 2010). Thus, GAL-1 expression under a condition of viral infection may aid viral survival in the infected cells.

In the present study, we demonstrated that NS5B interacts with GAL-1 and RPS4X mRNA not only *in vitro* but also *in vivo* through the 3′-UTR portions. Further studies are required to elucidate how it is that NS5B-mRNA interactions modulate the expression of these genes.

In conclusion, we have identified a population of liver cellular mRNA that binds to the NS5B protein through the 3′-UTR regions by employing an iterative selection cycle using an mRNA library enriched with the 3′-UTR portions. Elucidating the interaction of the cellular mRNA 3′-UTRs and HCV NS5B RdRp will be a great help for understanding the mechanism by which HCV replicates in the intracellular environment in association with cellular factors.

Materials and methods

Purification of the recombinant NS5B protein

A recombinant glutathione *S*-transferase (GST)–NS5B (strain BK, genotype 1b) fusion protein that lacks the C-terminal 21 amino acids portion was produced and purified as described (Uchiyama et al., 2002). Purification of the GST-free NS5B protein was performed as described in our previous work (Kanamori et al., 2009). The protein concentration was measured by Coomassie staining or the Bradford method.

Preparation of a 3′-UTR cDNA library

Total RNA was extracted from HepG2 cells with TRIzol (Invitrogen) and partially digested with 0.01 U/ml RNase T1 (Ambion) at room temperature for 15 min. The mRNA 3′-UTR fragment was obtained using the PolyATract mRNA isolation system (Roche) and reverse transcribed with a cDNA library construction kit (Takara). Subsequently, cDNA was size fractionated by polyacrylamide gel electrophoresis to obtain the 3′-cDNA fragments sizing from 75 to 150 base pairs. The recovered cDNA fragments were digested by EcoRI and NotI, and subcloned into the pAP3neo pre-digested vector (Takara). Plasmids containing cDNA clones were linearized by NotI and an RNA pool was synthesized (pool 0) using T7 RNA polymerase.

Selection procedure

Approximately 10^5 individual RNA molecules were used as the initial RNA pool (pool 0) for selecting the cellular RNA 3′-UTRs that bind tightly to the HCV NS5B protein. A 100 μg of the GST-NS5B protein was incubated with 200 μl of glutathione agarose beads (Sigma) in 2 ml of LG buffer (2 mM Tris–HCl [pH 7.5], 1 mM EDTA, 5 mM DTT, 0.5% Triton X-100, 0.2 mM PMSF and 40% glycerol) containing 150 mM NaCl for 1 h at 4 °C. Following incubation, the beads were washed 5 times with LG buffer containing 150 mM NaCl to remove any unbound protein, and subsequently washed 5 times with binding buffer (8 mM Hepes [pH 7.9], 40 mM NaCl, 2 mM EDTA, 0.2 mM DTT, 0.2 mM PMSF and 1.6% glycerol). The initial pool of RNA was added to the beads and allowed to bind for 15 min at 30 °C in 800 μl of binding buffer containing 50 μg/ml tRNA. After 5 washes with the binding buffer, the RNA molecules that remained bound to the beads were eluted by incubation with 300 μl of proteinase K digestion buffer (20 mM Tris–HCl [pH 7.9], 100 mM NaCl, 10 mM EDTA, 1% SDS and 400 μg/ml proteinase K) for 30 min at 37 °C. RNA was extracted with phenol/chloroform and precipitated with ethanol. The recovered RNA was reverse transcribed using primer A (5′-TAATACGACTCACTATAGGGAATTCCCGGG-3′) and PCR-amplified (25 cycles; 30 s at 96 °C, 30 s at 55 °C, 45 s at 72 °C) by adding primer B (5′-ACGTGCGGCCGCTTTTTTTTTT-3′). Primers were removed from the reaction mixture with a Micro Bio-Spin 30 Chromatography Column (Bio-Rad). The recovered PCR products were used as templates to synthesize an RNA pool (pool 1).

Selection cycles were repeated until we obtained the pool 8 RNA. For the selection cycles 2–7, 25 μg of GST-NS5B and 50 μl of glutathione agarose were used. For the selection cycles 3–7, RNA pools were first absorbed to GST-attached glutathione agarose beads to remove RNA molecules that were bound to the GST or agarose beads, and the unbound RNA was used for the selection process.

Filter binding assay and RNA gel mobility shift experiments

Filter binding assay was used as described previously to measure the binding strength of each RNA molecule to NS5B (Kanamori et al., 2009).

For RNA gel mobility shift analysis, synthetic RNA molecules were labeled with ^{32}P at the 5′-end using T4 polynucleotide kinase and [γ-^{32}P]ATP. The labeled RNA probes were purified by denaturing gel electrophoresis and subsequently incubated with proteins in a total of 10 μl of binding buffer containing 50 μg/ml tRNA for 20 min at 22 °C. The samples were loaded on a 6% polyacrylamide gel in a 1:4 dilution of Tris-borate/EDTA buffer and electrophoresed at 300 V and 4 °C. Gels were visualized with a Bioimage Analyzer BAS2000.

RdRp assay

For competition experiments, RdRp activity was measured by using the poly(C)-oligo(G) system, as previously described (Uchiyama et al., 2002). Reactions were performed in a 100-μl reaction mixture containing 20 mM Tris–HCl (pH 7.5), 25 mM KCl, 1 mM DTT, 5 mM $MnCl_2$, 0.01% BSA, 100 U/ml RNasin, 5 μg/ml actinomycin D, 0.4 μCi [^{3}H]GTP, 0.5 μg/ml poly(C), 50 nM oligo $(rG)_{12}$ and 80 nM purified NS5B in the presence (50 or 250 nM) or absence of competitor RNA molecules. After incubation for 2 h at room temperature, samples were precipitated with 3 volumes of ethanol in the presence of 10 μl of 3 M sodium acetate (pH 5.2) and 3 μl of ethachinmate (Nippon Gene), and transferred to GF/C glass microfiber filters (Whatman). The filters were rinsed with 70% ethanol and air dried. The filter bound radioactivity was measured using a scintillation counter. Experiments were performed in triplicate and the standard deviations are indicated in Fig. 4.

GAL-1 or RPS4X 3′-UTR mRNA was used as the template for RNA synthesis. RNA templates (final concentration of 125 nM) were incubated with 80 nM recombinant GST-NS5B in a 100-μl reaction mixture containing 50 mM Tris–HCl (pH 8), 50 mM NaCl, 100 mM KCl, 5 mM $MnCl_2$, 1 mM DTT, 20 μg/ml actinomycin D, 20 U RNasin, 10% glycerol, 0.5 mM each rATP, rCTP, rGTP and 5 μM rUTP with 40 μCi [α-^{32}P]UTP (10 mCi/ml) for 2 h at 25 °C. Reaction products were extracted with phenol/chloroform, precipitated with ethanol and 3 M sodium acetate, purified with Centri-Sep spin columns (Life Technologies), denatured in 95% formamide, and separated on a 6% polyacrylamide/8 M urea gel.

Immunoprecipitation-reverse transcription-PCR (IP-RT-PCR)

COS cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. RCYM1 is a Huh7 derived cell line and was a generous gift from Dr. Takeya Tsutsumi (The University of Tokyo) (Murakami et al., 2006). RCYM1 cells harbor a genome-length dicistronic HCV genotype 1b full-length RNA and were cultured in a monolayer in Dulbecco's modified Eagle's medium with 10% fetal bovine serum and 0.5 mg/ml G418 and harvested at the exponential growth phase.

We constructed a FLAG-NS5B expressing plasmid, pFLAG-5B to express the NS5B protein in COS cells. The NS5B coding region cDNA lacking the C-terminal 21 amino acid portion was PCR amplified using the primer C (5′-AAGAATGCGGCCGCGTCAATGTCCTACACA-3′) and primer D (5′-CCGGAATTCTTAACGGGGTCGGGCACG-3′), with the pGEX-4T-2-5bC21 plasmid as the template (Uchiyama et al., 2002). Following digestion with EcoRI and NotI, the DNA fragment was ligated into the pFLAG-CMV-4 vector (Sigma) to obtain the pFLAG-5B plasmid. The pFLAG-5B expression plasmid was transiently transfected into COS cells using Lipofectamine™ LTX (Invitrogen). Cells were incubated 24 h at 37 °C in a CO_2 incubator before harvesting for the analysis.

Cytoplasmic extracts of transfected COS cells containing the FLAG-fused NS5B protein were collected using the FLAG® Immunoprecipitation Kit (Sigma) by following the manufacturer's instructions. Expression of the FLAG-NS5B fusion protein was