[36] Suganami T, Tanimoto-Koyama K, Nishida J, Itoh M, Yuan X, Mizuarai S, et al. Role of the Toll-like receptor 4/NF-kappaB pathway in saturated fatty acid-induced inflammatory changes in the interaction between adipocytes and macrophages. Arterioscler Thromb Vasc Biol 2007;27:84–91.
[37] Vinciguerra M, Veyrat-Durebex C, Moukil MA, Rubbia-Brandt L, Rohner-Jeanrenaud F, Foti M. PTEN down-regulation by unsaturated fatty acids triggers hepatic steatosis via an NF-kappaBp65/mTOR-dependent mechanism. Gastroenterology 2008;134:268–280.
[38] Vinciguerra M, Sgroi A, Veyrat-Durebex C, Rubbia-Brandt L, Buhler LH, Foti M. Unsaturated fatty acids inhibit the expression of tumor suppressor phosphatase and tensin homolog (PTEN) via microRNA-21 up-regulation in hepatocytes. Hepatology 2009;49:1176–1184.
[39] Hatanaka H, Tsukui M, Takada S, Kurashina K, Choi YL, Soda M, et al. Identification of transforming activity of free fatty acid receptor 2 by retroviral expression screening. Cancer Sci 2010;101:54–59.
[40] Kang S, Bader AG, Vogt PK. Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. Proc Natl Acad Sci USA 2005; 102:802–807.
[41] Pettinelli P, Videla LA. Up-regulation of PPAR-γ mRNA expression in the liver of obese patients: an additional reinforcing lipogenic mechanism to SREBP-1c induction. J Clin Endocrinol Metab 2011;96:1424–1430.
[42] Blanco-Aparicio C, Renner O, Leal JF, Carnero A. PTEN, more than the AKT pathway. Carcinogenesis 2007;28:1379–1386.
[43] Li G, Robinson GW, Lesche R, Martinez-Diaz H, Jiang Z, Rozengurt N, et al. Conditional loss of PTEN leads to precocious development and neoplasia in the mammary gland. Development 2002;129:4159–4170.
[44] Ackler S, Ahmad S, Tobias C, Johnson MD, Glazer RI. Delayed mammary gland involution in MMTV-AKT1 transgenic mice. Oncogene 2002;21: 198–206.
[45] Hutchinson J, Jin J, Cardiff RD, Woodgett JR, Muller WJ. Activation of Akt (protein kinase B) in mammary epithelium provides a critical cell survival signal required for tumor progression. Mol Cell Biol 2001;21:2203–2212.
[46] Tamura M, Gu J, Matsumoto K, Aota S, Parsons R, Yamada KM. Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. Science 1998;280:1614–1617.
[47] Jia S, Liu Z, Zhang S, Liu P, Zhang L, Lee SH, et al. Essential roles of PI(3)K-p110beta in cell growth, metabolism and tumorigenesis. Nature 2008;454: 776–779.
[48] Day CP, James OF. Steatohepatitis: a tale of two "hits"? Gastroenterology 1998;114:842–845.
[49] Malaguarnera L, Madeddu R, Palio E, Arena N, Malaguarnera M. Heme oxygenase-1 levels and oxidative stress-related parameters in non-alcoholic fatty liver disease patients. J Hepatol 2005;42:585–591.

nature COMMUNICATIONS

ARTICLE

Received 30 Mar 2011 | Accepted 11 May 2011 | Published 7 Jun 2011

DOI:10.1038/ncomms1345

MicroRNA122 is a key regulator of α-fetoprotein expression and influences the aggressiveness of hepatocellular carcinoma

Kentaro Kojima[1,*], Akemi Takata[1,*], Charles Vadnais[2,*], Motoyuki Otsuka[1], Takeshi Yoshikawa[1], Masao Akanuma[3], Yuji Kondo[1], Young Jun Kang[4], Takahiro Kishikawa[1], Naoya Kato[5], Zhifang Xie[6], Weiping J. Zhang[6], Haruhiko Yoshida[1], Masao Omata[1], Alain Nepveu[2] & Kazuhiko Koike[1]

α-fetoprotein (AFP) is not only a widely used biomarker in hepatocellular carcinoma (HCC) surveillance, but is also clinically recognized as linked with aggressive tumour behaviour. Here we show that deregulation of microRNA122, a liver-specific microRNA, is a cause of both AFP elevation and a more biologically aggressive phenotype in HCC. We identify CUX1, a direct target of microRNA122, as a common central mediator of these two effects. Using liver tissues from transgenic mice in which microRNA122 is functionally silenced, an orthotopic xenograft tumour model, and human clinical samples, we further demonstrate that a microRNA122/CUX1/microRNA214/ZBTB20 pathway regulates AFP expression. We also show that the microRNA122/CUX1/RhoA pathway regulates the aggressive characteristics of tumours. We conclude that microRNA122 and associated signalling proteins may represent viable therapeutic targets, and that serum AFP levels in HCC patients may be a surrogate marker for deregulated intracellular microRNA122 signalling pathways in HCC tissues.

[1] Department of Gastroenterology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan. [2] Goodman Cancer Center and Departments of Oncology, Biochemistry and Medicine, McGill University, Montreal, Quebec H3A 1A3, Canada. [3] Division of Gastroenterology, The Institute for Adult Diseases, Asahi Life Foundation, Tokyo 100-0005, Japan. [4] Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, California 92037, USA. [5] Unit of Disease Control Genome Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan. [6] Department of Pathophysiology, Second Military Medical University, Shanghai 200433, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M. Otsuka (email: otsukamo-tky@umin.ac.jp).

The incidence of hepatocellular carcinoma (HCC), the third most common cause of cancer-related mortality worldwide[1], is currently increasing[2]. The recent discovery of the efficacy of sorafenib, a multikinase inhibitor, as a treatment for patients with advanced HCC, has represented a major breakthrough in the clinical management, although the survival benefit has been shown to be less than 3 months[3]. No other effective therapy is currently available for patients with advanced disease[4]. As such, there is a continuing need to develop novel therapeutics and approaches for treatment of advanced HCC[5].

To develop targeted cancer therapies, we must first identify aberrantly regulated molecular pathways specific to this cancer. Clinically, it is also important to discover useful and convenient surrogate serum biomarkers that reflect aberrations in molecular pathways due to the molecular mechanisms of their expression, to identify the deregulated intracellular signalling pathways and to spare the patients from invasive clinical tests.

Currently, α-fetoprotein (AFP) is the most widely used serum biomarker for HCC surveillance[6]. Although the regulation of AFP gene expression is not fully understood, p53 (ref. 7), β-catenin[8] and the recently identified zinc-finger protein, ZBTB20 (ref. 9), have been reported to be involved. Furthermore, whereas mounting clinical evidence indicates that AFP elevation is linked to a more aggressive tumour phenotype characterized by vascular invasion, metastasis and poor differentiation[10,11], it remains to be determined whether the two phenotypes represent anything more than coincidental epiphenomena[12].

MicroRNAs (miRNAs) are short, single-stranded, non-coding RNAs. Although first identified in *Caenorhabditis elegans*[13], miRNAs are now known to be expressed in most organisms, from plants to vertebrates[14]. Primary miRNAs, which possess stem-loop structures, are processed into mature miRNAs by Drosha and Dicer RNA polymerase III. These mature miRNAs then associate with the RNA-induced silencing complex, and the resulting complex directly binds to the 3′-untranslated regions of target messenger RNAs to act as suppressors of translation and gene expression. Thus, depending on the target mRNAs, miRNAs are responsible for the control of various biological functions including cell proliferation, apoptosis, differentiation, metabolism, oncogenesis and oncogenic suppression[15–17].

MiRNA122 (miR122) is a tissue-specific miRNA that is most abundant in the liver[18], wherein it is responsible for the maintenance of fatty acid metabolism[19,20] and circadian rhythms[21]. As shown for other tissue-specific miRNAs[22], expression of miR122 has been reported to be downregulated in carcinomas, particularly in more malignant tumours, although these results remain controversial because of conflicting reports[23–26]. The biological significance of the downregulation of miR122 expression in HCC at the molecular level has not yet been fully elucidated.

In the present study, we explored the role of microR122 in HCC by silencing it both in human HCC cells and in a transgenic mouse model. Our molecular analysis enabled us to define the complex regulatory cascades underlying the clinically recognized link between raised AFP levels and a more aggressive phenotype in HCC.

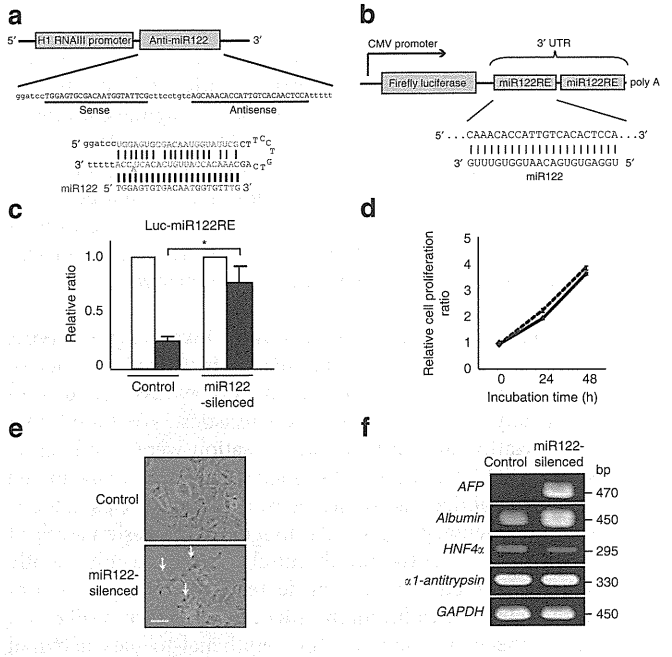


Figure 1 | Establishment of miR122-silenced HCC cell lines. (**a**), The miRZip122 construct, which yields a functional single-stranded full-length antisense miRNA complementary o miR122, processed from a stem-loop-stem molecule and transcribed from the constitutively active H1 RNA III promoter. Several mismatched nucleotides were included to efficiently produce a single-stranded antisense miR122 molecule. (**b**) Schematic representation of the luciferase reporter used to examine miR122 activity. The firefly luciferase gene, driven by a cytomegalovirus (CMV) promoter, contains two tandem miR122-binding sites (miR122-responsive elements; miR122-RE) in its 3′-UTR. (**c**) The suppressive effects of miR122 precursor expression (black bar) on luciferase activity in control and miR122-silenced cells. Test values were normalized to those obtained from the cells transfected with a miRNA precursor-non-expressing empty vector (white bar), which were set to 1. Data represent the mean±s.d. of three independent experiments using Huh7 cells. $*P<0.05$ (*t*-test). Similar results were obtained using PLC/PRF/5 cells. (**d**) Control (solid line) and miR122-silenced cells (dashed line) were plated at a density of 2×10^4 cells per well. After incubation for 24 and 48 h, numbers of cells were calculated as described in the Methods. Data represent the mean±s.d. of three independent experiments using Huh7 cells. Similar results were obtained using PLC/PRF/5 cells. (**e**) Changes in cellular morphology of Huh7 miR122-silenced cells are shown. Arrows indicate pseudopodia. Scale bar, 50 μm. Similar phenotypes were observed using PLC/PRF/5 cells. (**f**) The expression of several hepatocyte markers in control and Huh7 miR122-silenced cells was evaluated by semi-quantitative RT-PCR. Representative results from three independent experiments using Huh7 cells are shown. Similar results were obtained using PLC/PRF/5 cells.

Results

Establishment of miR122-silenced HCC cell lines. To characterize the functional consequences of miR122 downregulation in HCC cells, Huh7 and PLC/PRF/5 cells were stably transduced with a lentivirus that expresses RNA hairpins that produce mature antisense RNA designed to silence miR122 function. These cells were selected on the basis of their relatively high levels of miR122 expression[24,27]. Several mismatches were intentionally inserted into the RNA hairpin sequences to produce more stable templates for miR122 binding and sequestration and to perturb the participation of miR122 in RNA-induced silencing complex-associated inhibition of translation (Fig. 1a).

To confirm effective miR122 silencing in transduced cells, we analysed the luciferase activity of reporters containing miR122-binding sites (the function of which is suppressed by miR122 overexpression) (Fig. 1b) in miR122-silenced and control cells. As expected, overexpression of the miR122 precursor greatly suppressed luciferase activity in control cells (Fig. 1c). In contrast, this suppressive effect was significantly reduced in miR122-silenced cells (Fig. 1c), indicating that miR122 was indeed functionally silenced.

To characterize the biological changes that result from the loss of miR122 function, we next analysed cell proliferation, morphology and differentiation in miR122-silenced cells. The rates of cell

proliferation were comparable between control and silenced cells (Fig. 1d); however, miR122-silenced cells exhibited a larger number of distinct pseudopodia (Fig. 1e). Next, as miR122 is specifically expressed in the liver, we hypothesized that it may have a role in hepatocyte differentiation and, therefore, we investigated the expression of several hepatocyte markers by semi-quantitative RT–PCR. We observed an increase in AFP expression and a slight elevation of albumin expression in miR122-silenced cells, but the expression levels of other hepatocyte markers, such as hepatocyte nuclear factor 4α (HNF4α) and α_1-antitrypsin, did not change (Fig. 1f).

MiR122-silenced HCC cells exhibit a more invasive phenotype. Because miR122-silenced cells exhibited an increased number of pseudopodia, we next characterized phenotypes associated with more biologically aggressive cell characteristics. We found that actin polymerization and pseudopod formation were significantly increased in miR122-silenced cells (Fig. 2a). The increase in the number of pseudopodia was confirmed by a quantitative pseudopodia assay (Supplementary Fig. S1). Although the expression levels of the mesenchymal marker α-smooth muscle actin were only slightly increased, we observed a significant decrease in the expression of the epithelial marker E-cadherin in miR122-silenced cells (Fig. 2b). Furthermore, the expression of other epithelial-to-mesenchymal transition markers such as fibronectin, N-cadherin, snail and Zeb1 was altered in miR122-silenced cells (Supplementary Fig. S1). These findings are consistent with the notion that loss of miR122 function leads to a more malignant phenotype.

We next performed scratch and invasion assays to characterize the invasive phenotype of miR122-silenced cells. Rates of cell migration and of cell invasion were significantly increased in miR122-silenced cells (Fig. 2c,d). As the proliferation rates of control and miR122-silenced cells were similar (Fig. 1d), altogether these results suggest that inhibition of miR122 function in HCC cells may lead to increases in malignancy-related cellular properties.

To investigate the molecular mechanisms underlying these cellular phenotypes, we assessed the activity of RhoA and Rac1, which are small GTPases that are closely associated with cell migration and invasion[28]. Although Rac1 activity did not significantly change, RhoA activity significantly increased in miR122-silenced cells (Fig. 2e), suggesting that the increase in cell migration and invasion in miR122-silenced cells may result from increased RhoA activity.

AFP expression is increased in miR122-silenced HCC cells. As we observed an increase in AFP expression in miR122-silenced cells (Fig. 1f), we next sought to quantify AFP concentrations in culture supernatants using an enzyme-linked immunosorbent assay (ELISA). AFP levels were approximately three times higher in the supernatant of miR122-silenced cells as compared with control cells (Fig. 3a). Consistent with this observation, immunofluorescence staining for AFP produced a stronger cytoplasmic signal and quantitative RT–PCR revealed a tenfold increase in AFP mRNA levels in miR122-silenced cells (Fig. 3b,c).

The 3′-UTR of the AFP mRNA did not contain predicted miR122 target sequences, based on sequence analyses performed using miRNA target search engines such as TargetScan (http://www.targetscan.org), suggesting that it is unlikely that miR122 directly regulates AFP expression. Therefore, to characterize the mechanisms underlying increased AFP expression in miR122-silenced cells, we first assessed the stability of the AFP mRNA in miR122-silenced cells. As expected, AFP mRNA stability was unaffected by silencing of miR122, as the amount of mRNA was comparable between control and miR122-silenced cells at 6, 12 and 24 h after inhibition of new transcription by treatment with actinomycin D (Fig. 3d). The increase in AFP mRNA levels in the absence of changes in mRNA stability suggested that transcription of AFP was increased

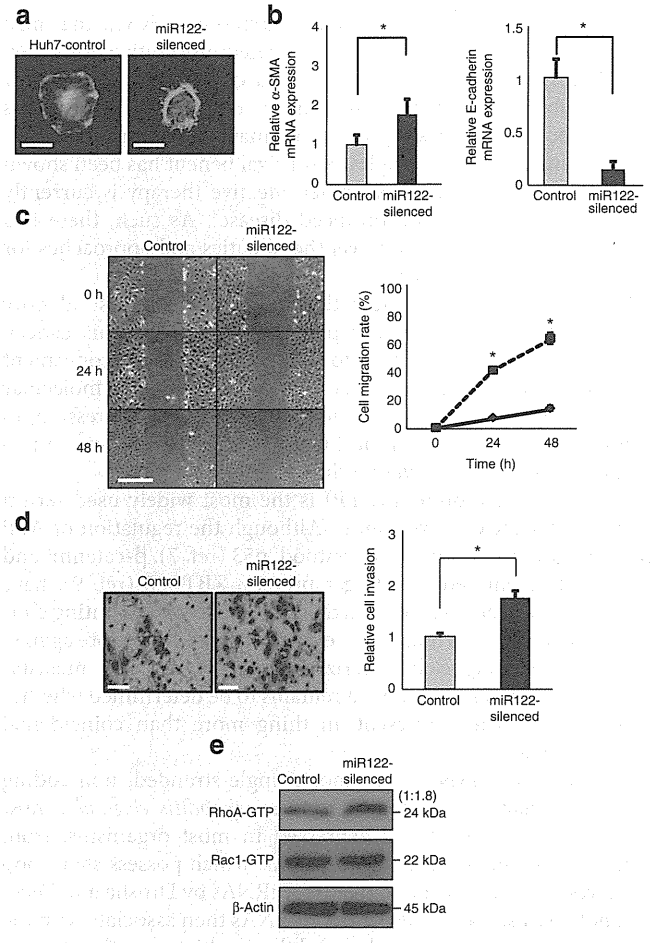


Figure 2 | HCC cells silenced for miR122 function exhibit a more invasive phenotype. (**a**) Cells were treated with 2 ng ml^{-1} TGF-β for 12 h, and actin filaments were stained with Alexa Fluor 488-conjugated phalloidin. Representative results from two independent experiments using Huh7 cells are shown. Similar results were obtained using PLC/PRF/5 cells. Scale bar, 50 μm. (**b**) Expression levels of α−smooth muscle actin and E-cadherin mRNAs were assessed by quantitative RT–PCR. Values shown represent mRNA expression levels in experimental cells relative to control cells. Data represent the mean±s.d. of three independent experiments using Huh7 cells. *$P<0.05$ (*t*-test). Similar results were obtained for PLC/PRF/5 cells. (**c**) The degree of cell migration was characterized using a scratch assay. The ratio of migrating cells was significantly increased in miR122-silenced cells at 24 and 48 h after scratching. Left panels show representative images. Right panel shows the results from cell counts for four randomly chosen fields per experiment. Data are represented as the mean±s.d. of three experiments using Huh7 control (solid line) and miR122-silenced cells (dashed line). *$P<0.001$ (*t*-test). Similar results were obtained for PLC/PRF/5 cells. (**d**) The degree of cell invasion was examined using cell invasion chambers. Representative images of stained invaded cells (left). The relative cell invasion ratio after normalization to control cell invasion levels (right). Data represent the mean±s.d. of three independent experiments using Huh7 cells. Scale bar, 100 μm. *$P<0.01$ (*t*-test). Similar results were obtained for PLC/PRF/5 cells. (**e**) Rho and Rac1 activity was determined by comparing the amounts of active GTP-bound RhoA (RhoA-GTP) and Rac1 (Rac1-GTP) between control and miR122-silenced cells. The indicated numbers represent the relative expression levels. Equal loading in pull-down assays was confirmed by analysis of β-actin levels in the cell lysates. Representative results of five independent experiments using Huh7 cells are shown. Similar results were obtained using PLC/PRF/5 cells.

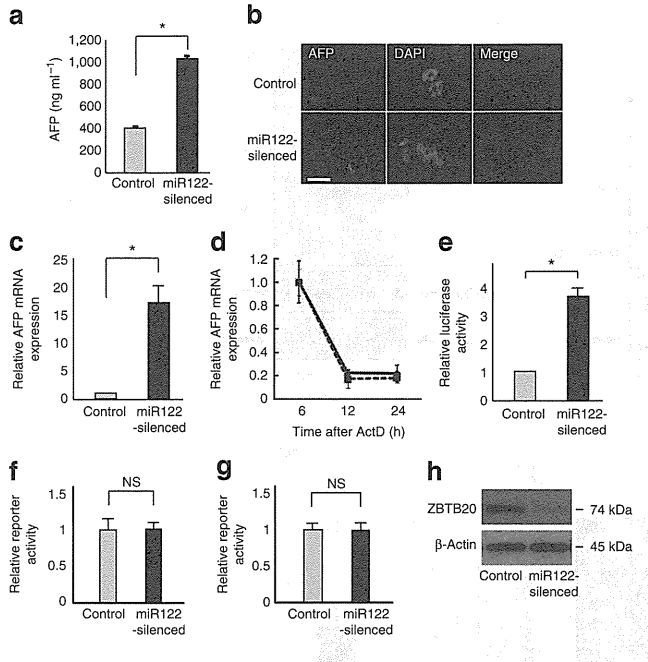


Figure 3 | Increased expression of AFP in HCC cells silenced for miR122 function. (**a**) The AFP concentration in the culture medium was determined by ELISA. Data represent the mean±s.d. of three independent experiments using Huh7 cells. *$P<0.01$ (*t*-test). Similar results were obtained for PLC/PRF/5 cells. (**b**) Immunofluorescent staining for AFP in the cytoplasm of control and miR122-silenced Huh7cells. Representative images of stained cells from three independent experiments are shown. Scale bar, 50 μm. (**c**) Amounts of AFP mRNA in Huh7 control and miR122-silenced cells were determined by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). Similar results were obtained using PLC/PRF/5 cells. (**d**) AFP mRNA stability in Huh7 control (solid line) and miR122-silenced (dashed line) cells was determined by quantitative RT-PCR at 6, 12, and 24 h after treating cells with 10 μl ml^{-1} actinomycin D. Data represent the mean±s.d. of three independent experiments. Similar results were observed using PLC/PRF/5 cells. (**e**) AFP promoter activity was measured in reporter assays using Huh7 cells and AFP promoter-luciferase construct. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). Similar results were obtained using PLC/PRF/5 cells. (**f**, **g**) Luciferase assays were performed using reporter plasmids to measure p53 (**f**) and β-catenin (**g**) activities. Data represent the mean±s.d. of three independent experiments. Similar results were obtained using PLC/PRF/5 cells. (**h**) ZBTB20 protein levels in miR122-silenced cells. A representative result from three independent experiments using Huh7 cells is shown. Similar results were obtained using PLC/PRF/5 cells.

in miR122-silenced cells as a result of increased AFP promoter activity. Indeed, AFP promoter activity was almost four times higher in miR122-silenced cells than in control cells, as assessed by a reporter assay (Fig. 3e).

Because AFP promoter activity is in part regulated by p53 (ref. 7), we assessed p53 activity using reporter constructs. However, no changes in p53 activity were detected in miR122-silenced cells (Fig. 3f). As mutation of β-catenin has also been reported to be involved in upregulation of AFP expression[8], we next analysed β-catenin activity in miR122-silenced cells. Similar to p53, no change in β-catenin activity was evident in miR122-silenced cells compared with control cells (Fig. 3g).

Recently, it was reported that ZBTB20 acts as a repressor of AFP transcription[9]. This result led us to assess the expression of the ZBTB20 protein in miR122-silenced cells. Indeed, western blot analysis revealed that ZBTB20 expression was decreased in miR122-silenced cells (Fig. 3h). However, as ZBT20 lacks the presence of predicted miR122 target sequences based on computational searches of the 3′-UTR, it was also unlikely that miR122 directly regulates ZBTB20 expression. These observations suggest that other indirect mechanisms may lead to decreased ZBTB20 expression in miR122-silenced cells.

CUX1 is the regulator of phenotypes in miR122-silenced cells. To explore the mechanisms by which miR122 regulates cell motility, invasion and AFP expression, we used computational searches to identify potential miR122 target genes with known functions related to these processes. This analysis led to the identification of Cut homeobox 1 (CUX1, also known as CCAAT-displacement protein/cut homeobox, CDP/Cux/Cut) through the presence of a high probability miR122 target site located in the 3′-UTR and a perfect match in the seed sequences. CUX1 is a transcription factor that regulates multiple processes including cell cycle progression, chromosomal segregation and cell migration[29,30]. Consistent with the effects of miR122 silencing described above, CUX1 was reported to modulate cell motility and invasion through the control of RhoA activity[31–33]. We observed that whereas CUX1 mRNA levels remained unchanged (Fig. 4a), there was a significant increase in the steady-state level of the CUX1 p200 and p110 isoforms in miR122-silenced cells (Fig. 4b).

To investigate the contribution of CUX1 upregulation to the increase in AFP expression and invasive properties observed in miR122-silenced cells, we knocked down CUX1 protein expression using lentiviruses expressing CUX1 short hairpin RNAs (shRNAs) (Fig. 4c). In the resulting double-knockdown cells, AFP protein expression in cell-culture supernatant and cell invasion were both reduced to levels similar to that of the parental Huh7 cells (Fig. 4d–f).

CUX1 represses ZBTB20 expression via miR214. We next assessed whether miR122 directly targets CUX1 by constructing a luciferase reporter construct that possessed a portion of the CUX1 3′-UTR containing the putative miR122 target site (Fig. 5a). Co-transfection experiments revealed that luciferase activity was suppressed by overexpression of a miR122 precursor-expressing plasmid (Fig. 5b). This suppressive effect was prevented by introducing two point mutations into the seed sequences of the miR122 target site (Fig. 5a,b), demonstrating that miR122 directly targets these sequences.

To confirm these effects, we generated 293T cell lines that stably expressed the miR122-precursor construct by transducing cells with miR122 precursor-expressing lentiviruses tagged with green fluorescent protein (Supplementary Fig. S2a). As expected, the anti-miR122 construct did not affect control 293T cells, owing to the lack of miR122 expression. However, the anti-miR122 construct greatly enhanced luciferase activity in 293T cells stably expressing the miR122-precursor, confirming that miR122 was transduced into the 293 T cells (Supplementary Fig. S2b). Consistent with the results described above, these cells exhibited decreased expression of CUX1, particularly the p200 isoform, and also showed a modest, but reproducible, increase in ZBTB20 expression (Fig. 5c). These results suggest that miR122 directly regulates CUX1 protein expression, which in turn may regulate ZBTB20 expression.

Because CUX1 can function as a transcriptional modulator[29], we initially hypothesized that CUX1 is a direct regulator of ZBTB20 transcription. However, quantitative RT–PCR analysis revealed that levels of the ZBTB20 mRNA were unchanged in miR122-silenced cells compared with controls (Fig. 5d). To explain the discrepancy between unchanged levels of ZBTB20 mRNA and decreases in protein expression levels in miR122-silenced cells, we searched for miRNAs that could potentially target the ZBTB20 3′-UTR. Based on computational searches, miR214 and miR375 were identified as candidate ZBTB20-regulatory miRNAs. Although levels of miR375 were unchanged in miR122-silenced cells (Fig. 5e), expression of miR214 was significantly increased (Fig. 5e).

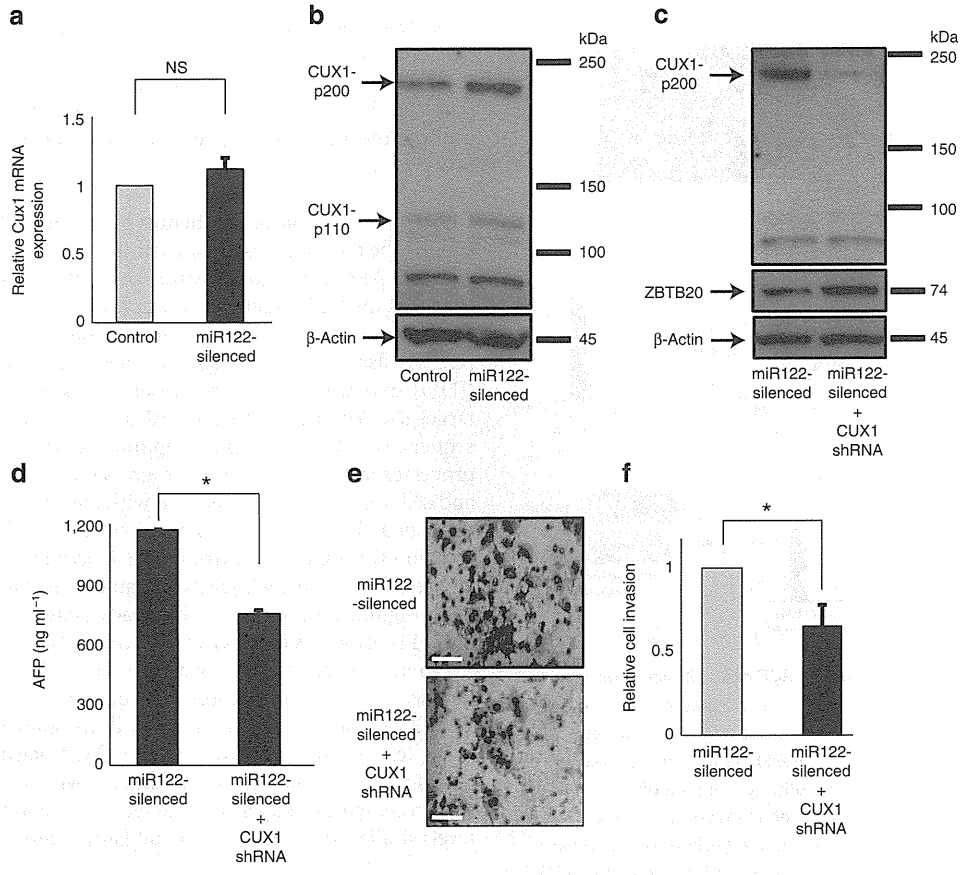


Figure 4 | CUX1-mediated regulation of AFP expression and phenotypic changes in miR122 functionally silenced cells. (**a**) CUX1 mRNA levels in control and miR122-silenced cells were analysed by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments using Huh7 cells. (**b**) p200 and p110 CUX1 protein levels were increased in miR122-silenced cells compared with control cells. A representative result from three independent experiments using Huh7 cells is shown. Similar results were obtained for PLC/PRF/5 cells. (**c**) CUX1 and ZBTB20 protein expression in double CUX1/miR122 knockdown Huh7 cells. Similar results were obtained using PLC/PRF/5 cells. (**d**) AFP concentrations in the culture medium supernatants were determined by ELISA. Data represent the mean±s.d. of three independent experiments. *$P<0.01$ (*t*-test). Similar results were observed using PLC/PRF/5 cells. (**e**, **f**) The change of cell invasion ratio by CUX1 knockdown in miR122-silenced Huh7 cells. Representative images of stained invading cells are shown (**e**). The relative cell invasion ratio after normalization to control invasion levels is shown (**f**). Data represent the mean±s.d. of three independent experiments. Scale bar, 100 μm. *$P<0.01$ (*t*-test). Similar results were obtained using PLC/PRF/5 cells.

To assess whether miR214 directly targeted the ZBTB20 3′-UTR, we constructed a luciferase reporter with the region of the ZBTB20 3′-UTR that contains the putative miR214 target site. Reporter assays revealed that luciferase activity was indeed suppressed by overexpression of the miR214 precursor, suggesting that miR214 directly targets the ZBTB20 3′-UTR and suppresses its expression (Fig. 5f). Consistent with these findings, cells that stably overexpressed the miR214 precursor exhibited decreased levels of ZBTB20 protein expression (Fig. 5g).

The putative promoter regions of miR214 contain multiple CUX1 binding sites as revealed by MATCH, a transcription factor binding site search engine (http://www.gene-regulation.com). A scanning chromatin immunoprecipitation (ChIP) experiment, followed by real-time PCR, using a series of primer pairs, confirmed that CUX1 binds to multiple genomic sites in the miR214 promoter region (Fig. 6a). We therefore hypothesized that CUX1 may regulate miR214 transcription. Consistent with this notion, we found that miR214 expression was decreased in CUX1 knockdown Huh7 cells (Fig. 6b). The role of CUX1 as an activator of miR214 transcription was further verified by knocking down or overexpressing CUX1 in another cell line. Levels of miR214 decreased following the constitutive knockdown of CUX1 with shRNA (Fig. 6c). In contrast, retroviral infection with a vector expressing p110 CUX1 led to an increase in miR214 (Fig. 6d). These findings were confirmed using doxycycline-inducible CUX1 shRNA. As previously observed for other transcriptional targets of CUX1 (refs 30,34), levels of miR214 were reduced in the presence of doxycycline, and then returned to levels higher than in untreated cells upon removal of the doxycycline inducer miR214 (Fig. 6e).

Next, to assess the contribution of miR214 to the control of ZBTB20 expression in miR122-silenced cells, we measured ZBTB20 expression after parallel silencing of miR214 in miR122-silenced cells. Although ZBTB20 protein expression was reduced by almost 50% by miR122 silencing, it was restored to >90% of control levels by miR214 silencing (Supplementary Fig. S3). Thus, CUX1-induced miR214 regulates, at least in part, ZBTB20 expression in miR122-silenced cells, leading to the upregulation of AFP expression.

Regulation of of CUX1 and AFP expression by miR122 was also confirmed in other HCC cell lines in which miR122 was overexpressed or silenced. Northern blotting showed that the expression of miR122 was relatively low in Hep3B and HepG2 cells, but was relatively high in Huh1, Huh7 and PLC/PRF/5 cells (Supplementary Fig. S4a). We therefore overexpressed the miR122 precursor in Hep3B and HepG2 cells and silenced miR122 in Huh1, Huh7 and PLC/PRF/5 cells (Supplementary Fig. S4b). CUX1 expression was respectively suppressed and enhanced by miR122 precursor overexpression and miR122 silencing (Supplementary Fig. S4c). In contrast, AFP expression was respectively enhanced and suppressed

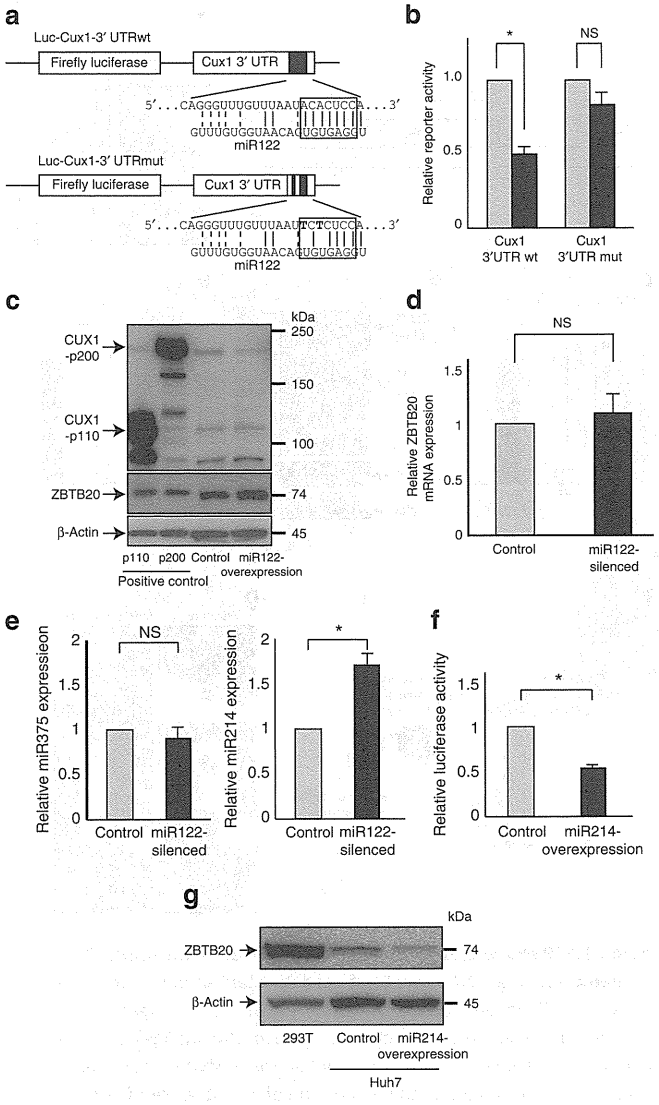


Figure 5 | MiR122 directly targets CUX1. (**a**) A luciferase reporter carrying a region of the wild type CUX1 3′UTR containing the putative miR122 target site (Luc-CUX1-3′UTRwt) was used to assess the effects of miR122 on expression of CUX1. A second luciferase reporter with two nucleotide mutations (indicated in bold) in the seed sequences (indicated by a rectangle) of the putative miR122 target sites (Luc-CUX1-3′UTRmut) was also utilized to assess specificity. (**b**) Huh7 cells were co-transfected with Luc-CUX1-3′UTRwt or Luc-CUX1-3′UTRmut and either an empty control vector (white bar) or a miR122 precursor expression plasmid (black bar). Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**c**) CUX1 and ZBTB20 expression in 293T cells-expressing the miR122 precursor. Cell lysates transiently transfected with CUX1 p200 or p110 expression plasmids were used as positive controls. Representative results from four independent experiments are shown. (**d**) ZBTB20 mRNA levels in miR122-silenced Huh7 cells were determined by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments. Similar results were obtained using PLC/PRF/5 cells. (**e**) Levels of miR375 (left) and miR214 (right) in miR122-silenced Huh7 cells were analysed by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**f**) Huh7 cells were co-transfected with Luc-ZBTB20-3′UTR and either an empty control vector or an miR214 precursor expression plasmid. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**g**) ZBTB20 expression was decreased in Huh7 miR214-overexpressing cells. 293T cell lysate was used as a positive control. Representative results from four independent experiments are shown.

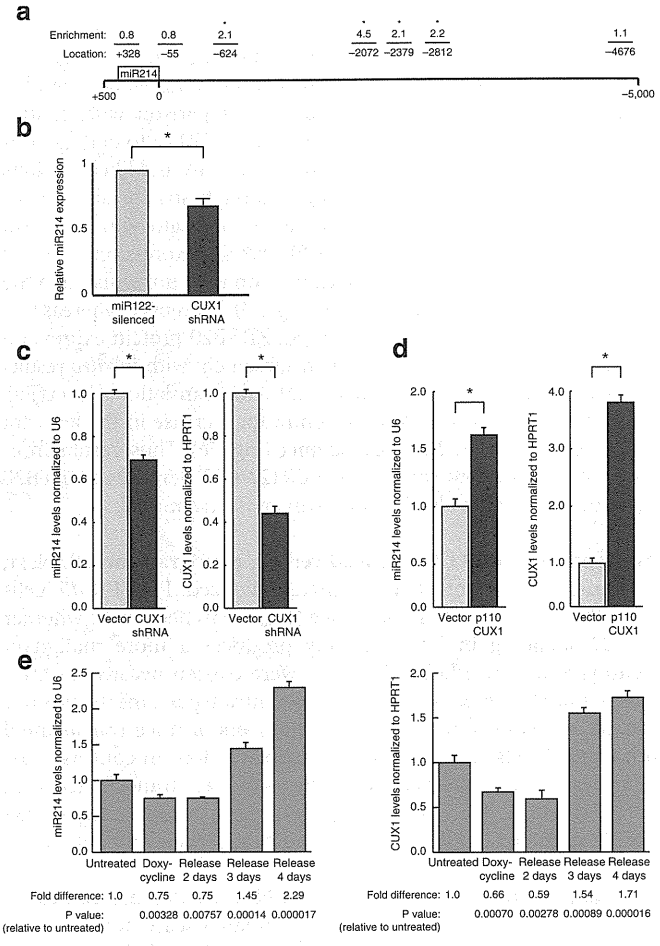


Figure 6 | CUX1 regulated miR-214 expression. (**a**), CUX1 enrichment at the miR214 locus. ChIP assays were performed using Hs578T cells. Fold enrichment and location of the center of each qPCR amplicon are shown. *$P<0.05$ (*t*-test). (**b**) MiR214 expression levels after CUX1 knockdown in miR122-silenced Huh7 cells were determined by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). Similar results were obtained using PLC/PRF/5 cells. (**c**) Levels of miR214 and CUX1 RNA in Hs578T cells infected with an empty vector or a lentiviral vector constitutively expressing CUX1 shRNA. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**d**) Hs578T cells were infected with a retrovirus expressing p110 CUX1. Levels of miR214 and CUX1 mRNA were measured 1 day later. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**e**) Levels of miR214 and CUX1 mRNA in Hs578T cells expressing a doxycycline-inducible CUX1 shRNA. Levels are shown prior to treatment, after 5 days of doxycycline treatment, and 2, 3 and 4 days after withdrawal of doxycycline. Fold changes with the mean±s.d. of three independent experiments and *P*-values are shown (*t*-test).

(Supplementary Fig. S4d), confirming that AFP expression is regulated by an miR122-CUX1 pathway in multiple HCC cell lines.

These results indicate that functional silencing of miR122 leads to an increase in CUX1 protein expression, resulting in repression of ZBTB20 through an increase in miR214 expression. Repression of ZBTB20, in turn, leads to an increase in AFP expression. Because CUX1 is a modulator of cell motility and invasion[35–37], upregulation of this protein also enhances RhoA activity, increasing the malignant properties of cancer cells.

Expression of CUX1-related molecules in miR122-silenced mice. To explore the pathway delineated above in an *in vivo* model, we gen-

erated transgenic mice expressing antisense miR122 under the control of an H1 promoter (Fig. 7a) to inhibit the function of endogenous miR122 (ref. 38). *In situ* hybridization analysis in these mice revealed weak miR122 staining in liver tissue in comparison with control mice, likely due to binding of the anti-sense miR122 to endogenous miR122, which produces a double-stranded DNA and likely inhibits hybridization of the probe (Fig. 7b). Although structural development of the liver appeared normal based on haematoxylin and eosin staining (Supplementary Fig. S5), AFP mRNA expression (Fig. 7c) and p200 and p110 CUX1 protein expression were upregulated in the liver of anti-miR122 transgenic mice (Fig. 7d). Moreover, whereas levels of ZBTB20 mRNA were unchanged, ZBTB20 protein expression was decreased in the liver (Fig. 7d), in agreement with *in vitro* results demonstrating the regulation of ZBTB20 at the translational level (Fig. 5c,d). This was associated with a significant increase in the levels of miR214 in anti-miR122 transgenic mice (Fig. 7e). Thus, results from mouse liver tissue confirm that the miR122/CUX1/miR214/ZBTB20 regulatory pathway is also functional in an *in vivo* model.

Invasiveness of miR122-silenced cells in xenograft model. Next, we transplanted control and miR122-silenced PLC/PRF/5 cells under the liver capsule of nude mice (Fig. 7f) to determine whether miR122 silencing in HCC actually produces a more malignant phenotype *in vivo*. PLC/PRF/5 cells were chosen because of their transplantability in nude mice[39]. Neither intrahepatic metastases nor vascular invasion were detected in the livers of mice transplanted with control cells at 4 weeks post-transplantation. In contrast, vascular invasion was observed in the livers of mice transplanted with miR122-silenced HCC cells (Fig. 7g). These results suggest that miR122 silencing in HCC leads to a more aggressive phenotype.

HCC staging and the expression of miR122-related molecules. To assess the relevance of these results to human disease, we examined miR122 and AFP expression in several clinical-grade human HCC samples. We analysed miR122 expression by *in situ* hybridization (Fig. 8a) and AFP expression by immunohistochemistry (Fig. 8b). Both AFP expression and malignancy grading were inversely correlated with miR122 expression levels (Fig. 8c,d). In addition, CUX1, miR214 and ZBTB20 expression was also correlated with miR122 expression, as determined using serial sections (Supplementary Fig. S6a, b and c). These results, together with our studies in tissue culture systems and a transgenic mouse model, suggest that a reduction in the expression of miR122 increases AFP expression via a miRNA122-CUX1-miRNA214-ZBTB20 pathway and that the development of more biologically aggressive forms of HCC occurs via a miRNA122-CUX1-RhoA pathway (Supplementary Fig. S7). The miRNA-mediated mechanism described in this report may explain the clinically known link between increased AFP levels and more biologically aggressive cell characteristics in HCC.

Discussion

High AFP levels have been clinically shown to be an unfavourable prognostic factor in HCC patients[40]. In this study, we demonstrate that reduced expression of miR122 in HCC cells contributes to elevated AFP expression and, subsequently, a more aggressive phenotype. These results provide a molecular framework that explains the reported link between elevated AFP levels and a poor clinical outcome in HCC patients.

Clinically, high AFP expression is correlated with more biologically aggressive properties of HCC, as patients with high AFP levels have a significantly higher frequency of portal vein invasion and intrahepatic metastases. Additionally, these patients display significantly lower rates of recurrence-free survival and a trend towards lower overall survival[41]. In the present study, we have presented several lines of evidence indicating that decreased expression of miR122 in HCC leads to the two phenomena that

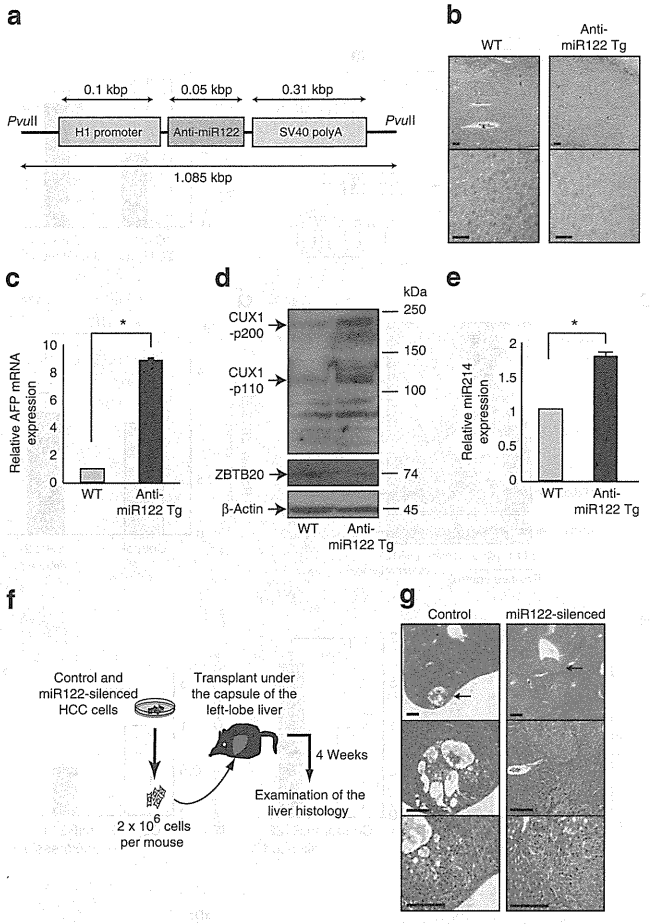


Figure 7 | AFP expression is increased in the liver of anti-miR122 transgenic mice. (**a**) DNA construct used to establishment transgenic mice in which miR122 is functionally silenced (anti-miR122 transgenic mice). This construct (like the construct used in the *in vitro* studies) generates a functional single-stranded full-length antisense miRNA complementary to miR122. (**b**) Confirmation of the expression of an antisense RNA directed against miR122 in anti-miR122 transgenic (anti-miR122 Tg) mice. Amounts of miR122 detected by *in situ* hybridization (blue/purple staining) in the liver tissues of anti-miR122 transgenic mice and WT mice. Results shown are representative of three independent experiments performed using littermates from four different mouse lines. Scale bar, 100 μm. No specific staining was observed when a negative control probe (LNA-scramble) was used. (**c**) AFP mRNA expression in mouse liver tissues was analysed by quantitative RT-PCR. Data represent the mean±s.d. of results for five mice in each group. *$P<0.05$ (*t*-test). (**d**) CUX1 and ZBTB20 expression in mouse liver tissues was assessed by Western blotting. A representative result from three independent experiments is shown. (**e**) Levels of miR214 in liver tissues were determined by quantitative RT-PCR. Data represent the mean±s.d. of three independent experiments. *$P<0.05$ (*t*-test). (**f**) Protocols of the orthotopic xenograft models of HCC cells. Control and miR122-silenced PLC/PRF/5 cells were prepared and injected under the capsule of the left-lobe of the nude mouse liver. Six mice were included in each group. At 4 weeks after transplantation, liver tissues were collected and sliced in series. Histological examination by H&E staining was performed to examine the tumour cell invasion status. (**g**) Representative liver histology images at 4 weeks after transplantation of tumour cells are shown. Whereas only the transplanted HCC cells beneath the capsule of the liver edge could be detected (arrow, upper left panel), neither intrahepatic metastasis nor vascular invasion was observed in the livers of mice transplanted with control cells. In contrast, vascular invasion by multiple tumour cells (arrow, upper right panel) was observed in mice transplanted with miR122-silenced cells. Scale bar, 500 μm. WT, wild type.

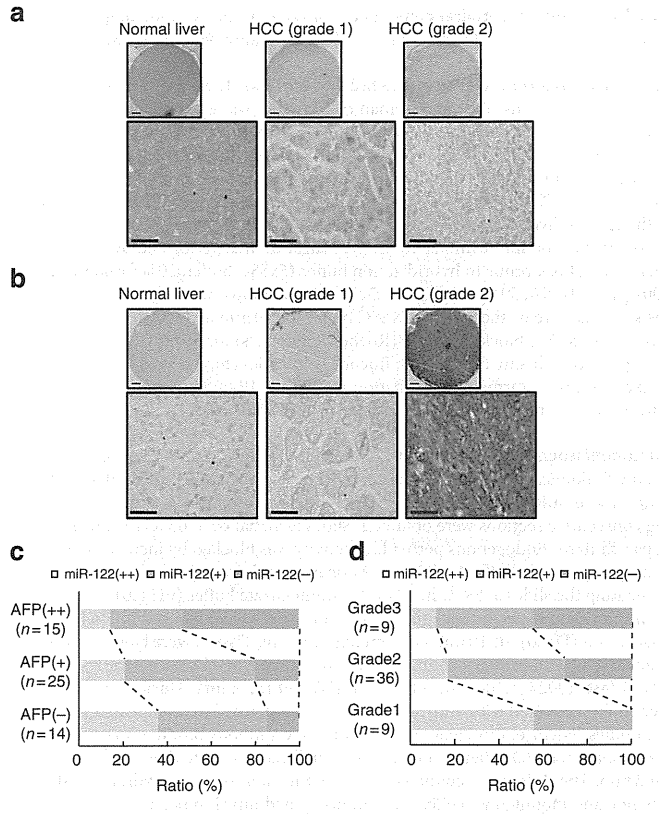


Figure 8 | AFP expression and HCC grade are inversely correlated with miR122 expression in human HCC samples. (**a**) Expression of miR122 in human clinical samples was assessed by *in situ* hybridization. Expression of miR122 (blue/purple staining) in grade 2 (more malignant) HCC samples was less than that of normal human liver tissues or grade 1 (less malignant) HCC samples. Representative images are shown. Nuclei were stained with FastRed. Scale bar, 500 μm. (**b**) AFP expression, shown in brown, was analysed by immunohistochemistry. AFP expression was higher in grade 2 HCC samples than in normal human liver tissue or grade 1 HCC samples in most cases. Representative images are shown. Scale bar, 500 μm. (**c**) Graph shows the correlation between miR122 and AFP expression. Increased AFP expression was correlated with decreased miR122 expression. (**d**) Graph shows the correlation between miR122 expression and malignancy grading of HCC. Increases in malignancy grading were correlated with decreases in miR122 expression.

are frequently observed simultaneously in the clinic: elevated expression of AFP and a more malignant biological phenotype. First, elevated AFP expression and greater cellular invasiveness were observed in miR122-knockdown cells *in vitro* and *in vivo*. Second, CUX1, which is linked with invasive characteristics in carcinoma cells[32,36,37], was shown to be involved in regulation of AFP expression and was identified as a direct target of miR122. Third, in human tissue samples from HCC patients, inverse correlations were observed between miR122 expression and AFP expression, and between miR122 expression and tumour grade. These data suggest that it is unlikely that the clinical correlation between elevated AFP levels and a more biologically aggressive phenotype in HCC is a coincidental epiphenomenon, but, instead provide a possible molecular explanation for the decrease in miR122 expression in HCC cells.

A recent study on liver development reported that liver-enriched transcription factors activate the expression of miR122, which in turn was found to promote terminal differentiation of hepatocytes through the silencing of CUX1 (ref. 42). In the present study, we confirmed that CUX1 is a direct target of miR122 and, in contrast to the situation in normal development, we showed that in grade 2 HCCs the decrease in miR122 is associated with higher CUX1 expression. High CUX1 expression was previously shown to inversely correlate with relapse-free and overall survival in high-grade breast cancers[36]. In transgenic mice, CUX1 was reported to cause various cancer-associated disorders depending on the specific isoform and tissue type expression[34,43–46]. In particular, expression of CUX1 caused organomegaly in several organs including the liver[43]. Hepatomegaly was associated with progression of lesions beginning with inflammation and leading to the development of mixed cell foci, hyperplasia and even HCCs, although in this last case statistical significance was not achieved because of the small size of the transgenic cohort[47]. The underlying mechanisms for the role of CUX1 in cancer is complex and is likely to involve both cell-autonomous and non-cell autonomous effects. However, from cell-based assays it is clear that CUX1 has a role in at least three distinct processes: cell motility, cell cycle progression and chromosome segregation[30,31,36,48]. The knockdown of CUX1 using siRNA was shown to delay entry into S phase and to hinder cell motility and invasion[31,36,48]. In contrast, overexpression of p110 CUX1 was able to accelerate S phase entry and to stimulate proliferation, migration and invasion[35,48]. Moreover, CUX1 was shown to promote genomic instability following cytokinesis failure[30]

Regulation of AFP gene expression is a complex process mediated by a number of transcriptional activators and repressors that bind the AFP gene[7,8]. ZBTB20 was recently identified as a potent repressor of AFP transcription in knockout mouse studies[9]. Our results demonstrate that decreased miR122 expression leads to concomitant decreases in ZBTB20 protein expression. This effect is mediated through upregulation of CUX1, as CUX1 silencing in miR122-silenced cells was shown to lead to both recovery of ZBTB20 levels and reduced AFP expression. Furthermore, the increased expression level of ZBTB20 in CUX1 knockdown cells suggests that ZBTB20 expression is regulated by CUX1. This miR122/CUX1/miR214/ZBTB20/AFP pathway may explain the deregulated AFP expression observed in HCC cells. Additionally, the ability of CUX1 to activate RhoA and to regulate the expression of many proteins involved in cell motility may explain the increased migration and invasiveness associated with malignancy of HCC[31–36]. It should be noted that, although this analysis revealed a trend toward inverse correlation between expression of miR122 and expression of AFP, this correlation could not be applied to all cases examined. Therefore, the possibility of additional pathways that regulate AFP expression cannot be discounted. Nonetheless, our results demonstrate that a decrease in miR122 function is a key factor that contributes to the regulation of AFP expression in HCC.

MiR122 is the most abundant miRNA in the normal adult liver, comprising approximately 80% of all miRNAs[18]. The numerous reported roles of miR122 include regulation of cholesterol biosynthesis[19,20], hepatitis C virus replication[49] and maintenance of the adult liver phenotype[21]. Specific miRNAs are often involved in the differentiation of specific cells and tissues[50]. As miR122 is liver-specific, we reasoned that this miRNA may have a role in the differentiation of normal hepatocytes. In our study, transgenic mice in which miR122 was functionally silenced were found to exhibit elevated AFP levels, but did not display abnormal morphological development in the liver (at least, not up to the age of 12 weeks), suggesting that decreased miR122 expression itself does not cause cells to become transformed. Ongoing characterization of these mice will be required to fully determine the physiological roles of miR122 in the noncancerous liver *in vivo*.

In summary, we have shown that decreased miR122 expression in HCC is linked both to more biologically aggressive tumour behaviour and elevated AFP expression. Furthermore, both of these effects were shown to be mediated by increased expression of CUX1, a direct target of miR122. Similar strategies could also be used to develop new therapeutics and diagnostics for other cancers in which miRNAs that regulate both tumour characteristics and serum markers have been identified.

Methods

Cell culture. The human HCC cell lines, Huh7, PLC/PRF/5, HepG2, Hep3B and Huh1 were obtained from the Japanese Collection of Research Bioresources. The human embryonic kidney cell line, 293T and the human breast cancer cell line HS578T were obtained from the American Type Culture Collection . All cells were maintained in Dulbecco's modified Eagle medium, supplemented with 10% fetal bovine serum.

Mouse experiments. All experiments were carried out in compliance with the regulations of the Animal Use Committee of The University of Tokyo and The Institute for Adult Disease, Asahi Life Foundation.

Generation of transgenic mice in which miR122 was functionally silenced. Mice in which miR122 function was knocked down were generated using previously described protocols[38,51]. Briefly, a DNA fragment of 1,085 bp, containing the H1 promoter region, the coding region for the antisense miR122 stem-loop-stem RNA precursor, and a transcriptional terminator of five thymidines, was resected from the miRZip-122 construct described above by digestion with *Pvu*II. Proper silencing function of the resulting DNA was confirmed via transient transfection-based reporter assays that showed efficient knockdown of miR122 function. Stable C57BL/6 embryonic stem cell lines were generated by electroporation of the linearized transgene, and the resulting cells were injected into blastocysts by the UNITECH Company. Genotyping was performed by PCR using DNA isolated from tail snips. Four different mouse lines were maintained and the male littermates were used in experiments.

Chromatin immunoprecipitation assay. ChIP for CUX1 was performed as previously described[52]. For the scanning ChIP of the miR214 locus, realtime PCR analysis was performed using primer pairs specific for different regions of the promoters. Templates for the PCR reactions were 0.1% total input DNA (I), nonspecific DNA from sepharose beads alone (S), or chipped chromatin. The respective fold enrichment of the different DNA fragments are indicated relative to the DNA obtained by purification on sepharose beads without IgG (S). Enrichment was calculated using the HPRT locus as a reference.

Doxycyclin-induced shRNA against CUX1 system. For conditional knockdown of CUX1 in Hs578t cells, we took advantage of the Addgene plasmid 11643. HS578T cells were infected with pLVCT shCUX1(5,326–5,348)-tTRKRAB lentivirus as described[53]. At 48 h after infection, cells were split and cultured with or without doxycyclin at a final concentration of 2.5 $\mu g\,ml^{-1}$. Cells were used for experiments after 5 days of treatment. Doxycyclin was then removed from the culture media and cells were maintained for 4 days following release.

Cell proliferation assay. Relative cell proliferation was assessed using a Cell Counting Kit-8 (Dojindo Laboratories), as described previously[54].

Enzyme-linked immunosorbent assay. AFP levels in the cell-culture supernatant were examined using an AFP-specific ELISA kit supplied by an outsourcing company, SRL.

Western blot analysis. Protein lysates were prepared from cells or mouse liver for immunoblot analysis. Proteins were separated by SDS–polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes. After blocking with 5% dry milk to decrease nonspecific binding, membranes were probed with the appropriate primary antibodies. Primary antibodies were obtained from Abcam (ZBTB20, #ab48889, 1:500) and Santa Cruz Biotechnology (CDP, #sc-13024, 1:1,000). CUX1 antibodies (#861, 1:1,000) were generated as described previously[52]. Horseradish peroxidase-conjugated secondary antibodies were used to detect primary antibodies. Bound antibodies were detected using ECL Plus Western blotting detection reagents (GE Healthcare Life Sciences).

Scratch assay. The effects of miR122 knockdown on cellular migratory function were determined by evaluating cellular migration after scratching of a confluent monolayer of cells. Monolayers were cultured on 10 $\mu g\,ml^{-1}$ fibronectin-coated dishes and were scratched using a 200-µl pipette. Migration was analysed at the indicated time points after scratching.

***In vitro* invasion assay**. The effect of miR122 knockdown on invasive function was determined using BD BioCoat Matrigel Invasion Chambers (Becton Dickinson) according to the manufacturer's recommended protocol. Cell invasion was induced by removing the serum in the upper chamber. The number of invading cells was analysed after 22-h incubation. Cell numbers were counted in four randomly chosen fields at each time point.

Quantitative pseudopodia assay. Pseudopodium quantitation was performed using a Quantitative Pseudopodia Assay Kit (Chemicon) according to the manufacturer's instructions. Briefly, the upper chamber was coated with fibronectin and seeded with cells in serum-free medium. Serum was added to the lower chambers. 8 h later, pseudopodia on the lower surface were stained and eluted, and the absorbances of solubilized samples at 600 nm was measured using a microplate reader.

CUX1-knockdown lentiviral construct. Lentiviral particles expressing CUX1 shRNA were purchased from Santa Cruz Biotechnology (#sc-35051-V).

***In situ* hybridization to assess miR122 and miR214**. The expression of miR122 and miR214 in mouse liver and human HCC tissues was examined by *in situ* hybridization[55,56]. Locked nucleic acid (LNA)-scramble (negative control) and LNA-anti-miR122 and LNA-anti-miR214 probes were obtained from EXIQON. After deparaffinization, tissue sections were treated with 10 $\mu g\,ml^{-1}$ proteinase K for 5 min at 37 °C and refixed with 4% paraformaldehyde, followed by acetylation with 0.25% anhydrous acetic acid in 0.1 M Tris–HCl buffer (pH 8.0). Following pre-hybridization for 30 min at 48 °C, hybridization was performed overnight with each 20 nM LNA probe in hybridization buffer (5xSSC buffer, 50% formamide, 500 $\mu g\,ml^{-1}$ tRNA, 50 $\mu g\,ml^{-1}$ Cot-1 DNA). After completion of hybridization, the sections were washed with 0.1×SSC buffer for 10 min at 52 °C three times and blocked with DIG blocking buffer (Roche Diagnostics) for 30 min. Sections were then probed with anti-DIG (1:500; Roche Diagnostics) for 1 h at room temperature. Detection was performed by incubation in NBT/BCIP buffer (Promega) overnight. Nuclei were stained with Nuclear FastRed (Sigma-Aldrich).

Immunohistochemistry. Tissue arrays containing HCC tissues were purchased from US Biomax. To determine the correlations between AFP, ZBTB20, CUX1, miR122 and miR214 expression and HCC differentiation grade, slides carrying consecutive sections were obtained. Slides were baked at 65 °C for 1 h and deparaffinized. Endogenous peroxidase activity was blocked by incubation in 3% hydrogen peroxide buffer for 30 min. Antigen retrieval was performed by incubating the slides at 89 °C in 10 mM sodium citrate buffer (pH 6.0) for 30 min. To minimize nonspecific background staining, slides were blocked in 5% normal goat serum (Dako) for 10 min at room temperature. Tissues were labelled overnight at 4 °C with primary antibodies raised against AFP (Dako, #N1501, 1:100), CUX1 (#sc-13024, 1:100) and ZBTB20 (HPA016815, Sigma-Aldrich, 1:100). Slides were then incubated with anti-rabbit horseradish peroxidase-conjugated secondary antibodies (Nichirei Bioscience) for 1 h. Primary antibody binding was visualized by incubation in 3,3′-diaminobenzidine in buffered substrate (Nichirei Bioscience) for 5 min. The slides were counterstained with haematoxylin, dehydrated with ethanol, and mounted with Clarion mounting medium (Biomeda).

GTP-binding RhoA and Rac1 immunoprecipitation assay. The amount of RhoA activity was examined using an Active Rho Pull-down and Detection Kit (Thermo Fisher Scientific) according to the manufacturer's recommended protocol. The amount of GTP-bound RhoA protein (the active form of RhoA) was detected by Western blotting with the provided anti-RhoA antibody (1:100). Rac1 activity was similarly determined by using PAK-GST Protein Beads (#PAK02, Cytoskeleton) for pulldowns and anti-Rac1 antibodies (1:100) for subsequent Western blotting (#89856D, Thermo Fisher Scientific).

Orthotopic xenograft tumour model of HCC. Male BALB/c (nu/nu) nude mice were purchased from CREA Japan (Tokyo, Japan). The transplantation of tumour cells into mouse livers was performed using previously reported methods[57,58]. Briefly, 2×10^6 control or miR122-silenced PLC/PRF5 cells were suspended in 30 µl of PBS containing 1% Matrigel (Becton Dickinson). After anaesthesia, the liver was exposed through a surgical incision. Cells were slowly injected under the capsule of left lobe of the liver using a 28-gauge needle. When successful, a transparent bleb of cells could be seen through the liver capsule. After injection, light pressure was applied to the injection site with sterile gauze for 2 min to prevent bleeding and tumour cell leakage. The abdomen was then closed with sutures. Transplantation was successful in a total of 12 mice (6/group). At 4 weeks post-transplantation, liver tissues were collected, serially sectioned, and stained with haematoxylin and eosin.

Statistical analysis. Statistically significant differences between groups were determined using Student's *t*-test when variances were equal. When variances were unequal, Welch's *t*-test was used instead. *P*-values less than 0.05 were considered statistically significant.

Plasmid and stable cell line construction, reporter assays, RT–PCR, northern blotting and immunocytochemistry are described in the Supplementary Methods. All primer information is provided in Supplementary Table S1.

References

1. Parkin, D., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **55,** 74–108 (2005).
2. El-Serag, H. Epidemiology of hepatocellular carcinoma in USA. *Hepatol. Res.* **37,** S88–94 (2007).
3. Llovet, J. *et al.* Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* **359,** 378–390 (2008).
4. Greten, T. *et al.* Survival rate in patients with hepatocellular carcinoma: a retrospective analysis of 389 patients. *Br. J. Cancer* **92,** 1862–1868 (2005).
5. Greten, T., Korangy, F., Manns, M. & Malek, N. Molecular therapy for the treatment of hepatocellular carcinoma. *Br. J. Cancer* **100,** 19–23 (2009).
6. Di Bisceglie, A. Issues in screening and surveillance for hepatocellular carcinoma. *Gastroenterology* **127,** S104–S107 (2004).

7. Ogden, S. *et al. p53* targets chromatin structure alteration to repress alpha-fetoprotein gene expression. *J. Biol. Chem.* **276,** 42057–42062 (2001).
8. Peng, S. *et al.* High alpha-fetoprotein level correlates with high stage, early recurrence and poor prognosis of hepatocellular carcinoma: significance of hepatitis virus infection, age, p53 and beta-catenin mutations. *Int. J. Cancer.* **112,** 44–50 (2004).
9. Xie, Z. *et al.* Zinc finger protein ZBTB20 is a key repressor of alpha-fetoprotein gene transcription in liver. *Proc. Natl Acad. Sci. USA* **105,** 10859–10864 (2008).
10. Oishi, K. *et al.* Clinicopathologic features of poorly differentiated hepatocellular carcinoma. *J. Surg. Oncol.* **95,** 311–316 (2007).
11. Yamamoto, K. *et al.* AFP, AFP-L3, DCP, and GP73 as markers for monitoring treatment response and recurrence and as surrogate markers of clinicopathological variables of HCC. *J. Gastroenterol.* **45,** 1272–1282 (2010).
12. Matsumoto, Y. *et al.* Clinical classification of hepatoma in Japan according to serial changes in serum alpha-fetoprotein levels. *Cancer* **49,** 354–360 (1982).
13. Lee, R., Feinbaum, R. & Ambros, V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75,** 843–854 (1993).
14. Carrington, J. & Ambros, V. Role of microRNAs in plant and animal development. *Science* **301,** 336–338 (2003).
15. Bartel, D. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297 (2004).
16. Ambros, V. The functions of animal microRNAs. *Nature* **431,** 350–355 (2004).
17. Lu, J. *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435,** 834–838 (2005).
18. Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129,** 1401–1414 (2007).
19. Krützfeldt, J. *et al.* Silencing of microRNAs *in vivo* with 'antagomirs'. *Nature* **438,** 685–689 (2005).
20. Esau, C. *et al.* miR-122 regulation of lipid metabolism revealed by *in vivo* antisense targeting. *Cell Metab.* **3,** 87–98 (2006).
21. Gatfield, D. *et al.* Integration of microRNA miR-122 in hepatic circadian gene expression. *Genes Dev.* **23,** 1313–1326 (2009).
22. Yan, D. *et al.* MicroRNA-1/206 targets c-Met and inhibits rhabdomyosarcoma development. *J. Biol. Chem.* **284,** 29596–29604 (2009).
23. Kutay, H. *et al.* Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. *J. Cell. Biochem.* **99,** 671–678 (2006).
24. Coulouarn, C., Factor, V., Andersen, J., Durkin, M. & Thorgeirsson, S. Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. *Oncogene* **28,** 3526–3536 (2009).
25. Tsai, W. *et al.* MicroRNA-122 a tumor suppressor microRNA that regulates intra-hepatic metastasis of hepatocellular carcinoma. *Hepatology* **49,** 1571–1582 (2009).
26. Varnholt, H. *et al.* MicroRNA gene expression profile of hepatitis C virus-associated hepatocellular carcinoma. *Hepatology* **47,** 1223–1232 (2008).
27. Wong, Q. *et al.* MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1. *Gastroenterology* **135,** 257–269 (2008).
28. Sahai, E. & Marshall, C. RHO-GTPases and cancer. *Nat. Rev. Cancer* **2,** 133–142 (2002).
29. Sansregret, L. & Nepveu, A. The multiple roles of CUX1: insights from mouse models and cell-based assays. *Gene* **412,** 84–94 (2008).
30. Sansregret, L. *et al.* Cut homeobox 1 causes chromosomal instability by promoting bipolar division after cytokinesis failure. *Proc. Natl Acad. Sci. USA* **108,** 1949–1954 (2011).
31. Kedinger, V. & Nepveu1, A. The roles of CUX1 homeodomain proteins in the establishment of a transcriptional program required for cell migration and invasion. *Cell Adh. Migr.* **4,** 348–352 (2010).
32. Michl, P., Knobel, B. & Downward, J. CUTL1 is phosphorylated by protein kinase A, modulating its effects on cell proliferation and motility. *J. Biol. Chem.* **281,** 15138–15144 (2006).
33. Seguin, L. *et al.* CUX1 and E2F1 regulate coordinated expression of the mitotic complex genes Ect2, MgcRacGAP, and MKLP1 in S phase. *Mol. Cell Biol.* **29,** 570–581 (2009).
34. Kedinger, V. *et al.* p110 CUX1 homeodomain protein stimulates cell migration and invasion in part through a regulatory cascade culminating in the repression of E-cadherin and occludin. *J. Biol. Chem.* **284,** 27701–27711 (2009).
35. Michl, P. *et al.* CUTL1 is a target of TGF(beta) signaling that enhances cancer cell motility and invasiveness. *Cancer Cell* **7,** 521–532 (2005).
36. Aleksic, T. *et al.* CUTL1 promotes tumor cell migration by decreasing proteasome-mediated Src degradation. *Oncogene* **26,** 5939–5949 (2007).
37. Kunath, T. *et al.* Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. *Nat. Biotechnol.* **21,** 559–561 (2003).
38. Shouval, D. *et al.* Tumorigenicity in nude mice of a human hepatoma cell line containing hepatitis B virus DNA. *Cancer Res.* **41,** 1342–1350 (1981).
39. Nomura, F., Ohnishi, K. & Tanabe, Y. Clinical features and prognosis of hepatocellular carcinoma with reference to serum alpha-fetoprotein levels. Analysis of 606 patients. *Cancer* **64,** 1700–1707 (1989).
40. Johnson, P., Melia, W., Palmer, M., Portmann, B. & Williams, R. Relationship between serum alpha-foetoprotein, cirrhosis and survival in hepatocellular carcinoma. *Br. J. Cancer* **44,** 502–505 (1981).
41. Xu, H. *et al.* Liver-enriched transcription factors regulate microRNA-122 that targets CUTL1 during liver development. *Hepatology* **52,** 1431–1442 (2010).
42. Ledford, A. W. *et al.* Deregulated expression of the homeobox gene Cux-1 in transgenic mice results in downregulation of p27(kip1) expression during nephrogenesis, glomerular abnormalities, and multiorgan hyperplasia. *Dev. Biol.* **245,** 157–171 (2002).
43. Brantley, J. G., Sharma, M., Alcalay, N. I. & Heuvel, G. B. Cux-1 transgenic mice develop glomerulosclerosis and interstitial fibrosis. *Kidney Int.* **63,** 1240–1248 (2003).
44. Cadieux, C. *et al.* Mouse mammary tumor virus p75 and p110 CUX1 transgenic mice develop mammary tumors of various histologic types. *Cancer Res.* **69,** 7188–7197 (2009).
45. Cadieux, C. *et al.* Polycystic kidneys caused by sustained expression of Cux1 isoform p75. *J. Biol. Chem.* **283,** 13817–13824 (2008).
46. Cadieux, C. *et al.* Transgenic mice expressing the p75 CCAAT-displacement protein/Cut homeobox isoform develop a myeloproliferative disease-like myeloid leukemia. *Cancer Res.* **66,** 9492–9501 (2006).
47. Vanden Heuvel, G. B. *et al.* Hepatomegaly in transgenic mice expressing the homeobox gene Cux-1. *Mol. Carcinog.* **43,** 18–30 (2005).
48. Sansregret, L. *et al.* The p110 isoform of the CDP/Cux transcription factor accelerates entry into S phase. *Mol. Cell Biol.* **26,** 2441–2455 (2006).
49. Jopling, C., Yi, M., Lancaster, A., Lemon, S. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309,** 1577–1581 (2005).
50. Taulli, R. *et al.* The muscle-specific microRNA miR-206 blocks human rhabdomyosarcoma growth in xenotransplanted mice by promoting myogenic differentiation. *J. Clin. Invest.* **119,** 2366–2378 (2009).
51. Zhou, Y. *et al.* Chimeric mouse tumor models reveal differences in pathway activation between ERBB family- and KRAS-dependent lung adenocarcinomas. *Nat. Biotechnol.* **28,** 71–78 (2010).
52. Harada, R. *et al.* Genome-wide location analysis and expression studies reveal a role for p110 CUX1 in the activation of DNA replication genes. *Nucleic Acids Res.* **36,** 189–202 (2008).
53. Szulc, J., Wiznerowicz, M., Sauvain, M. O., Trono, D. & Aebischer, P. A versatile tool for conditional gene expression and knockdown. *Nat. Methods* **3,** 109–116 (2006).
54. Maeda, S. *et al.* Ikappa B kinasebeta/nuclear factor-kappaB activation controls the development of liver metastasis by way of interleukin-6 expression. *Hepatology* **50,** 1851–1860 (2009).
55. Elmén, J. *et al.* LNA-mediated microRNA silencing in non-human primates. *Nature* **452,** 896–899 (2008).
56. Bai, S. *et al.* MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. *J. Biol. Chem.* **284,** 32015–32027 (2009).
57. Yao, X. *et al.* A novel orthotopic tumor model to study growth factors and oncogenes in hepatocarcinogenesis. *Clin. Cancer Res.* **9,** 2719–2726 (2003).
58. Kim, M. *et al.* Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nat. Protoc.* **4,** 1670–1680 (2009).

Acknowledgments

This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan (#22390058, #22590718, #17016016 and #20390204) (to M. Otsuka, Y. Kondo, M. Omata and K. Koike), by Health Sciences Research Grants of The Ministry of Health, Labour and Welfare of Japan (Research on Hepatitis) (to K. Koike), by grants from the Takeda Science Foundation, Astellas Foundation for Research on Metabolic Disorders, Senri Life Science Foundation, the Foundation for Promotion of Cancer Research and the Mochida Memorial Foundation for Medical and Pharmaceutical Research (to M. Otsuka), and by the grant 019389 from the Canadian Cancer Society (to A.N.).

Author contributions

K. Kojima, M. Otsuka and A.N. planned the research and wrote the paper. K. Kojima, A.T., C.V., T.Y., Y. Kondo, Y. Kang and Z.X. performed the majority of the experiments. M.A., N.K., W.Z. and A.N. contributed materials. T.K. and H.Y. supported several experiments. M. Omata and K. Koike supervised the entire project.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Kojima, K. *et al.* MiRNA122 is a key regulator of α-fetoprotein expression and influences the aggressiveness of hepatocellular carcinoma. *Nat. Commun.* 2:338 doi: 10.1038/ncomms1345 (2011).

nature genetics

Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma

Vinod Kumar[1,2], Naoya Kato[3], Yuji Urabe[1], Atsushi Takahashi[2], Ryosuke Muroyama[3], Naoya Hosono[2], Motoyuki Otsuka[4], Ryosuke Tateishi[4], Masao Omata[4], Hidewaki Nakagawa[2], Kazuhiko Koike[4], Naoyuki Kamatani[2], Michiaki Kubo[2], Yusuke Nakamura[1,2] & Koichi Matsuda[1]

To identify the genetic susceptibility factor(s) for hepatitis C virus–induced hepatocellular carcinoma (HCV-induced HCC), we conducted a genome-wide association study using 432,703 autosomal SNPs in 721 individuals with HCV-induced HCC (cases) and 2,890 HCV-negative controls of Japanese origin. Eight SNPs that showed possible association ($P < 1 \times 10^{-5}$) in the genome-wide association study were further genotyped in 673 cases and 2,596 controls. We found a previously unidentified locus in the 5′ flanking region of *MICA* on 6p21.33 (rs2596542, $P_{combined} = 4.21 \times 10^{-13}$, odds ratio = 1.39) to be strongly associated with HCV-induced HCC. Subsequent analyses using individuals with chronic hepatitis C (CHC) indicated that this SNP is not associated with CHC susceptibility ($P = 0.61$) but is significantly associated with progression from CHC to HCC ($P = 3.13 \times 10^{-8}$). We also found that the risk allele of rs2596542 was associated with lower soluble MICA protein levels in individuals with HCV-induced HCC ($P = 1.38 \times 10^{-13}$).

It is estimated that more than 170 million people are infected with HCV worldwide[1]. Persistent HCV infection causes CHC and, subsequently, fatal liver diseases such as liver cirrhosis and HCC. Therefore, the treatment of HCV carriers is an issue of global importance. HCC is the third most common cause of cancer-related deaths[2], and HCV infection accounts for 30–70% of the individuals with HCC[3,4]. HCV-induced HCC is a multistep and progressive liver disease in which disease progression may be influenced by both environmental and genetic risk factors. The impact of host genetic variation on progression to CHC after HCV exposure is well documented by recent genome-wide association studies (GWAS)[5–7]. However, no comprehensive analyses have been performed to explore the genetic basis of HCV-induced HCC. Therefore, we conducted a GWAS for HCV-induced HCC.

We genotyped the DNA of 721 individuals with HCV-induced HCC and 2,890 HCV-negative controls (**Supplementary Table 1**) from BioBank Japan[8]. After the initial standard SNP quality filters, we obtained genotyping results for 432,703 SNPs for association analysis. Because progression from CHC to liver cancer is strongly affected by age and gender[3], we performed a logistic regression analysis by including age and gender as covariates at all tested loci in our analyses. The genetic inflation factor (λ) was 1.03, indicating that there is no or little population stratification (**Supplementary Fig. 1**). Although no SNPs cleared the GWAS significance threshold ($P < 5 \times 10^{-8}$) at this stage, we identified eight independent loci showing possible association ($P < 1 \times 10^{-5}$; **Supplementary Fig. 2**).

In the replication stage, 673 cases from an independent HCC cohort from the University of Tokyo and 2,596 HCV-negative controls from BioBank Japan were genotyped at these eight SNPs. We observed a significant replication of association at rs2596542 on chromosome 6p21.33 ($P = 8.62 \times 10^{-9}$, odds ratio (OR) = 1.44, 95% confidence interval (CI) 1.27–1.63; **Table 1**), whereas the remaining seven SNPs failed to replicate the association (**Supplementary Table 2**). Furthermore, the combination analysis of the GWAS and replication study data at rs2596542 revealed a highly significant association in which the frequency of the risk allele A is higher in cases ($P = 4.21 \times 10^{-13}$, OR = 1.39; **Fig. 1** and **Table 1**) after the age and gender adjustment, without any heterogeneity ($P = 0.24$) between the two stages. To further investigate the impact of rs2596542 on the complex nature of the HCV-induced HCC phenotype, we genotyped 1,730 individuals with CHC who had not developed liver cirrhosis or HCC during their recruitment. As a result, rs2596542 was found to have no association with chronic hepatitis C susceptibility ($P = 0.61$) but was significantly associated with progression from CHC to HCC ($P = 3.13 \times 10^{-8}$, OR = 1.36; **Table 2**).

Because heavy alcohol consumption (>50 g per day) as well as poor response to interferon (IFN) treatment were shown to be the major risk factors for HCC among individuals with CHC[9], we evaluated the effect of alcohol consumption as a confounding factor and found that rs2596542 remained highly significant even after adjustment for this factor (non-HCV versus HCC, OR = 1.39, $P = 1.22 \times 10^{-11}$; CHC versus HCC, OR = 1.25, $P = 2.31 \times 10^{-4}$; **Supplementary Table 3**). The major genotypes of HCV can be determined by a serotyping

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. [2]Center for Genomic Medicine, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan. [3]Unit of Disease Control Genome Medicine, The Institute of Medical Science, University of Tokyo, Tokyo, Japan. [4]Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. Correspondence should be addressed to K.M. (koichima@ims.u-tokyo.ac.jp).

Received 20 July 2010; accepted 23 March 2011; published online 17 April 2011; doi:10.1038/ng.809

Table 1 Association results of rs2596542 in the GWAS, replication stage and combined analysis

SNP	Chr. (locus)	Stage	Case RAF	Control RAF	*P*	OR (95% CI)
rs2596542 (A/G)	6 (*MICA*)	GWAS[a]	0.388	0.331	4.50×10^{-6}	1.34 (1.16–1.53)
		Replication[a]	0.413	0.331	8.62×10^{-9}	1.44 (1.27–1.63)
		Combined[a]	0.400	0.331	4.21×10^{-13}	1.39 (1.27–1.52)
				MH test	7.76×10^{-12}	1.35 (1.24–1.47)

We analyzed 1,394 cases with HCC (721 in the GWAS and 673 in the replication) and 5,486 controls (2,890 in GWAS and 2,596 in replication). Chr., chromosome; RAF, risk allele frequency (allele A); OR, odds ratio for the minor allele calculated by considering the major allele as a reference; MH, Mantel-Haenszel.
[a]*P* values and ORs are adjusted for age and gender by logistic regression analysis under an additive model.

assay that is based on the type-specific antibodies produced by the infected host[10]. A subgroup analysis for HCV serotypes or history of IFN therapy indicated that this variation is associated with HCC susceptibility independently of HCV genotypes or treatment response (**Supplementary Fig. 3**). Consistent with this result, rs1051796, which had $r^2 = 0.7$ and $D' = 0.95$ with rs2596542, was not associated with IFN response ($P = 0.89$) according to previously published data in the Japanese population[11].

rs2596542 is located within the class I major histocompatibility complex (MHC) region. The human MHC region encompasses the complex and extended linkage disequilibrium (LD) structure[12,13]. Several HLA alleles and genes within MHC region have been implicated in HCV infection or clearance or in response to treatment[14–16]. Therefore, we searched the whole 7.5-Mb extended MHC region using GWAS data to test the possibility of other associated loci. We found a moderate association peak at rs9275572 ($P = 4.99 \times 10^{-5}$), which is located between *HLA-DQA* and *HLA-DQB* loci (**Supplementary Fig. 4**). Subsequent replication and combination analyses at rs9275572 indicated a significant association with HCV-induced HCC ($P = 9.38 \times 10^{-9}$, OR = 1.30; **Supplementary Table 4**). The multiple logistic regression analysis to control for alcohol consumption along with age and gender also indicated a significant association at rs9275572 ($P = 3.21 \times 10^{-8}$, OR = 1.29; **Supplementary Table 5**). However, rs2596542 was not in high LD with rs9275572 ($D' = 0.41$, $r^2 = 0.16$), and both SNPs remained associated with HCC even after conditional analysis on each other and had small reductions in their ORs upon conditioned analysis (OR = 1.23, $P = 4.43 \times 10^{-6}$ and OR = 1.17, $P = 0.00059$, respectively; **Supplementary Table 6**). A haplotype analysis between these two markers showed four possible haplotypes, with haplotype AA showing higher risk (with OR = 1.44) compared to the major haplotype GG (**Supplementary Table 7**). However, the OR for the risk haplotype was 1.32 with $P = 2.31 \times 10^{-10}$ after comparing against all observed haplotypes in the population (**Supplementary Table 7**), which is weaker than that of rs2596542 alone (OR = 1.39, $P = 4.21 \times 10^{-13}$). Hence, the impact of rs2596542 is much stronger than the haplotype of two SNPs, suggesting that rs2596542 is a principal genetic factor in this region. We also found that rs9275572 has a moderate association with CHC susceptibility as well as progression from CHC to HCC ($P = 0.03$ and $P = 2.58 \times 10^{-5}$, OR = 1.09 and OR = 1.29, respectively; **Supplementary Table 8**). Because *HLA-DQ* and *HLA-DR* alleles were shown to be associated with viral persistence and early liver disease among Japanese individuals[16], further study will be needed to confirm whether the association at rs9275572 is because of its LD with *HLA-DQ* or *DR* alleles.

In this regard, it is interesting to note that rs9275572 had a very strong expression quantitative trait locus effect on *HLA-DQB1* ($\log_{10}$ odds (LOD) ≥ 19.48) and *HLA-DRB4* alleles (LOD ≥ 26.88)[17]. Thus, it will be important to test the functional effect of the common haplotype (AA; **Supplementary Table 7**), which tags the risk alleles at these two SNPs.

Two SNPs, rs12979860 and rs8099917, at the *IL28B* locus were reported to be associated with spontaneous clearance of HCV virus[18] and response to pegylated IFN-α and ribavirin therapy[11], respectively. However, we found no association at rs12979860 and rs8099917 in our dataset (**Supplementary Table 9**). Because we used non-HCV control subjects rather than subjects who had cleared HCV infection spontaneously, and because only about 20% of the cases with HCC had been treated with IFN, our study may not be suitable to detect associations at the *IL28B* locus. In addition, the protective C allele at rs12979860 is nearly fixed throughout east Asia, with a frequency of more than 91% in the Japanese population as compared to 67% in European Americans[6], indicating a role for other factors in spontaneous clearance.

The top associated SNP, rs2596542, is located 4.7 kb upstream of *MICA*, the MHC class I polypeptide-related sequence A gene, and 41.7 kb downstream of the *HLA-B* gene (**Supplementary Fig. 5**). The regional association plot at the rs2596542 locus, made using genotype data from the GWAS (**Fig. 1**) and imputation analysis (**Supplementary Fig. 6**), revealed that all of the modestly associated SNPs are tightly

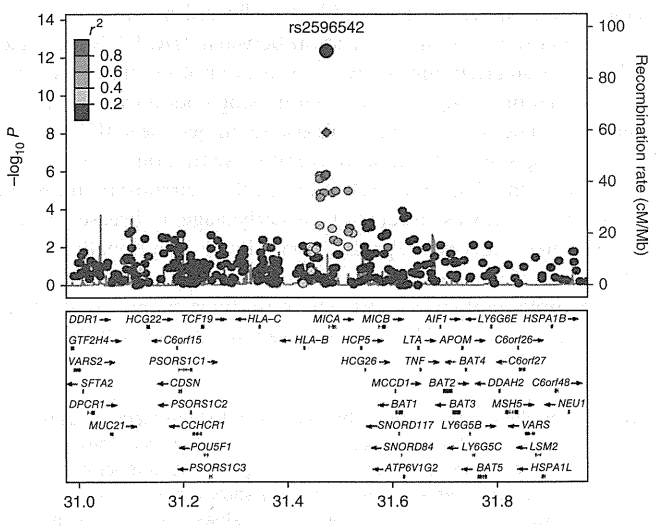


Figure 1 Regional association plot at rs2596542. Above, the *P* values of genotyped SNPs are plotted (as $-\log_{10}$ values) against their physical position on chromosome 6 (NCBI Build 36). The *P* value for rs2596542 at the GWAS stage, replication stage and combination analysis is represented by a purple diamond, circle and diamond, respectively. Estimated recombination rates from the HapMap JPT population show the local LD structure. Inset, the SNP's colors indicate LD with rs2596542 according to a scale from $r^2 = 0$ to $r^2 = 1$ based on pairwise r^2 values from HapMap JPT. Below, gene annotations from the UCSC genome browser.

Table 2 rs2596542 (A/G) is associated with progression from CHC to HCC

Subjects	RAF	(Comparison) *P*[a]	OR[a]	95% CI
Healthy	0.331			
CHC	0.333	(Healthy vs. CHC) 0.61	1.02	0.94–1.10
HCC	0.398	(CHC vs. HCC) 3.13×10^{-8}	1.36	1.22–1.51

We analyzed 5,486 controls, 1,730 CHC cases and 1,394 HCC cases. RAF, risk allele frequency (allele A); OR, odds ratio for the minor allele by considering the major allele as a reference.
[a]Calculated by logistic regression analysis, by PLINK upon age and gender adjustment under additive model.

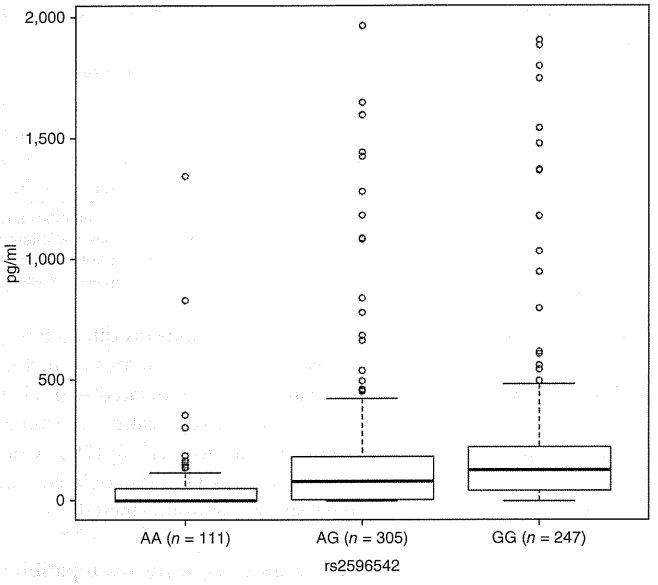
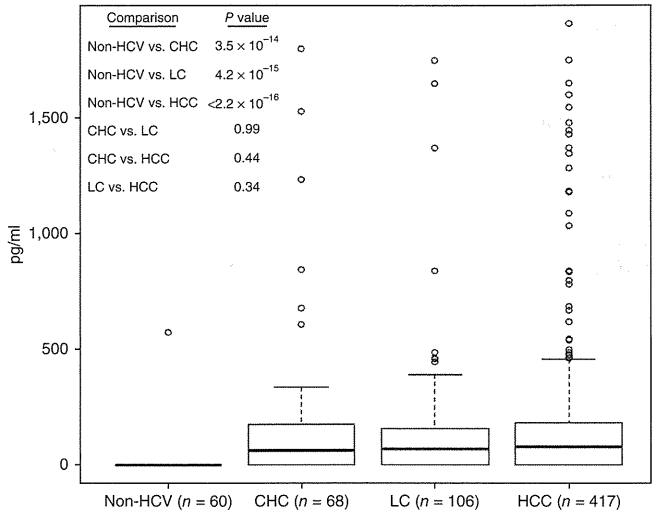


Figure 2 Correlation between soluble MICA levels and rs2596542 genotype. The *x* axis shows the genotypes at rs2596542, and the *y* axis shows the concentration of soluble MICA in pg/ml. The number of independent samples tested in each group is shown in parentheses. Each group is shown as a box plot, and the median values are shown as thick dark horizontal lines (median values of AA = 0, AG = 43.6 and GG = 77.74). The box covers the twenty-fifth to seventy-fifth percentiles, and the whiskers outside the box extend to the highest and lowest value within 1.5 times the interquartile range. Points outside the whiskers are outliers. We tested the difference in the median values among genotypes using the Kruskal-Wallis test ($P = 1.6 \times 10^{-13}$). We plotted the box plots using default settings in R (see URLs).

linked to rs2596542 ($r^2 > 0.4$) and are confined to the *MICA* gene locus. On the other hand, the imputation analysis of *HLA*-tagging SNPs did not show any evidence of linkage with rs2596542 (Online Methods and **Supplementary Table 10**), suggesting that *MICA* is a disease-associated candidate gene at this locus.

MICA is a membrane protein that acts as a ligand for NKG2D to activate anti-tumor effects through natural killer cells and $CD8^+$ T cells[19]. On the other hand, MICA is secreted into the serum by cleavage at the transmembrane domain with matrix metalloproteinases[20,21] and inhibits the anti-tumor effect of natural killer cells and $CD8^+$ T cells by blocking their action[22–24]. Elevated expression of both the membrane-bound and soluble forms of MICA (sMICA) have been reported in several cancers, including HCC[25–27]. Exon 5 of *MICA* encodes the transmembrane domain and contains a variable number of tandem repeats (VNTR) consisting of 4, 5, 6 or 9 repeats of GCT or one additional G nucleotide insertion into the 5-GCT-repeat allele (referred as A4, A5, A6, A9 and A5.1, respectively). The insertion of G (A5.1) causes a premature stop codon and subsequent loss of the transmembrane domain, leading to altered subcellular localization[28]. Therefore, we tested whether rs2596542 is in linkage with functional *MICA* VNTR alleles.

We further genotyped 673 cases with HCV-induced HCC and 890 non-HCV controls for the *MICA* VNTR locus with capillary-based electrophoresis (**Supplementary Fig. 7**). A case-control analysis revealed that the *MICA* VNTR is associated with HCV-induced HCC (global $P = 4.55 \times 10^{-7}$; **Supplementary Table 11**). Particularly, alleles A9 and A6 were associated with conferring a higher risk of HCC (OR = 1.73 and OR = 1.34, respectively), whereas the A5 and A5.1 alleles had a protective effect. Comparison of the genotypes at rs2596542 and the VNTR locus revealed that the A risk allele at rs2596542 is in LD with the A9 and A4 alleles, and the non-risk G allele is in LD with the A5 and A5.1 alleles, whereas we observed no linkage between an A6 allele and rs2596542 (**Supplementary Table 12**). We also genotyped 124 individuals with CHC; however, we observed no significant association between individuals with CHC and controls or individuals with CHC and HCC (**Supplementary Tables 13,14**).



We then tested whether the VNTR alleles, rs2596542 alleles, or VNTR-rs2596542 haplotypes had any association with MICA expression in individuals with HCV-induced HCC. We determined sMICA levels by ELISA using a total of 665 HCC serum samples (**Supplementary Table 15**). Notably, rs2596542 was significantly correlated with sMICA levels, and specifically, the risk genotype AA was associated with low levels of sMICA ($P = 1.38 \times 10^{-13}$; **Fig. 2**), whereas VNTR alleles (**Supplementary Fig. 8**) and VNTR- rs2596542 haplotypes (**Supplementary Table 16**) showed no strong association. The absence of any correlation between *MICA* VNTR alleles and sMICA suggests that sMICA levels are not regulated by post-translational processing or a premature stop codon caused by A5.1 alleles in individuals with HCC. We also examined the sMICA level in different stages of HCV-induced liver disease (in non-HCV subjects and those with CHC and HCV-induced liver cirrhosis) and found that sMICA level was elevated at the early stage of disease and was not correlated with disease progression (**Fig. 3**). Additionally, the risk allele A was also correlated with low sMICA levels in subjects with CHC (**Supplementary Fig. 9**). These findings suggest that MICA expression was induced by factors caused by chronic HCV infection,

Figure 3 Correlation between soluble MICA and HCV-related diseases. The *x* axis shows the disease stages after HCV infection, and the *y* axis shows the concentration of soluble MICA in pg/ml. The number of independent samples tested in each group is shown in parentheses. Each group is shown as a box plot, and the median values are shown as thick dark horizontal lines (median values of non-HCV = 0, CHC = 64.55, LC = 72.11 and HCC = 77.98). The box covers the twenty-fifth to seventy-fifth percentiles, and the whiskers outside the box extend to the highest and lowest value within 1.5 times the interquartile range. Points outside the whiskers are outliers. We tested the difference in the median values among the disease groups using the Wilcoxon rank test. The box plots were plotted using default settings in R. Non-HCV, individuals not exposed to HCV infection; CHC, individuals with chronic hepatitis C; LC, individuals with liver cirrhosis; HCC, individuals with hepatocellular carcinoma.

similar to various types of stresses such as viral infection, inflammation and heat shock[29,30]. The levels of sMICA were shown to be directly proportional to the level of membrane-bound MICA[25], and membrane bound MICA is essential for activating natural killer cells and $CD8^+$ T cells to eliminate virus-infected cells[19]. Considering the association of the risk allele A with low levels of sMICA, our findings suggest that the individuals who carry the rs2596542 A allele would express low levels of membrane-bound MICA in response to HCV infection, which thus leads to poor or no activation of natural killer cells and $CD8^+$ T cells against virus-infected cells. Eventually, these individuals are likely to progress from CHC to HCC. Notably, several SNPs that are in absolute linkage with rs2596542 are located within the promoter or enhancer region of *MICA* and may alter the binding of stress-inducible transcriptions factors such as heat shock proteins (**Supplementary Table 17**). In this regard, it is important to analyze the factors that regulate *MICA* expression, particularly in the context of CHC. Although, the molecular mechanism whereby *MICA* polymorphisms confer the risk of disease progression should be characterized in the future, our findings reveal a crucial role of genetic variations in the host innate immune system in the development of HCV-induced HCC.

URLs. R, http://cran.r-project.org/; PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/; Primer3 v0.3.0, http://frodo.wi.mit.edu/primer3/; LocusZoom, http://csg.sph.umich.edu/locuszoom/; FastSNP, http://fastsnp.ibms.sinica.edu.tw/pages/input_CandidateGeneSearch.jsp.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We would like to thank all the subjects and the members of the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan who donated their DNA for this work. We also thank the technical staff of the Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, and the Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo. This work was conducted as a part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government.

AUTHOR CONTRIBUTIONS

K.M. and Y.N. conceived of the study; Y.N., V.K., M.K. and K.M. designed the study; V.K., Y.U., R.M. and N.H. performed genotyping; V.K., Y.N. and K.M. wrote the manuscript; A.T. and N. Kamatani performed quality control at the genome-wide phase; Y.N., K.M., H.N. and M.K. managed DNA and serum samples belonging to BioBank Japan; N. Kato, R.T., M. Otsuka, M. Omata and K.K. managed replication DNA and serum samples; V.K. analyzed the data, performed VNTR genotyping, ELISA and summarized the whole results; Y.N. obtained funding for the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Global Burden of Hepatitis C Working Group. Global burden of disease (GBD) for hepatitis C. *J. Clin. Pharmacol.* **44**, 20–29 (2004).
2. Parkin, D.M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **55**, 74–108 (2005).
3. Umemura, T., Ichijo, T., Yoshizawa, K., Tanaka, E. & Kiyosawa, K. Epidemiology of hepatocellular carcinoma in Japan. *J. Gastroenterol.* **44** (Suppl 19), 102–107 (2009).
4. Vong, S. & Bell, B.P. Chronic liver disease mortality in the United States, 1990–1998. *Hepatology* **39**, 476–483 (2004).
5. Ge, D. *et al.* Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. *Nature* **461**, 399–401 (2009).
6. Thomas, D.L. *et al.* Genetic variation in *IL28B* and spontaneous clearance of hepatitis C virus. *Nature* **461**, 798–801 (2009).
7. Rauch, A. *et al.* Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* **138**, 1338–1345, 1345.e1–7 (2010).
8. Kamatani, Y. *et al.* A genome-wide association study identifies variants in the *HLA-DP* locus associated with chronic hepatitis B in Asians. *Nat. Genet.* **41**, 591–595 (2009).
9. Schütte, K., Bornschein, J. & Malfertheiner, P. Hepatocellular carcinoma–epidemiological trends and risk factors. *Dig. Dis.* **27**, 80–92 (2009).
10. Tanaka, T. *et al.* Significance of specific antibody assay for genotyping of hepatitis C virus. *Hepatology* **19**, 1347–1353 (1994).
11. Tanaka, Y. *et al.* Genome-wide association of *IL28B* with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat. Genet.* **41**, 1105–1109 (2009).
12. Anonymous. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* **401**, 921–923 (1999).
13. de Bakker, P.I. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* **38**, 1166–1172 (2006).
14. Kuniholm, M.H. *et al.* Specific human leukocyte antigen class I and II alleles associated with hepatitis C virus viremia. *Hepatology* **51**, 1514–1522 (2010).
15. Wang, J.H. *et al.* Ethnic and geographical differences in HLA associations with the outcome of hepatitis C virus infection. *Virol. J.* **6**, 46 (2009).
16. Singh, R., Kaul, R., Kaul, A. & Khan, K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. *World J. Gastroenterol.* **13**, 1770–1787 (2007).
17. Dixon, A.L. *et al.* A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007).
18. Ge, D. *et al.* Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. *Nature* **461**, 399–401 (2009).
19. Bauer, S. *et al.* Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. *Science* **285**, 727–729 (1999).
20. Salih, H.R., Rammensee, H. & Steinle, A. Cutting edge: down-regulation of MICA on human tumors by proteolytic shedding. *J. Immunol.* **169**, 4098–4102 (2002).
21. Waldhauer, I. *et al.* Tumor-associated MICA is shed by ADAM proteases. *Cancer Res.* **68**, 6368–6376 (2008).
22. Jinushi, M. *et al.* Impairment of natural killer cell and dendritic cell functions by the soluble form of MHC class I-related chain A in advanced human hepatocellular carcinomas. *J. Hepatol.* **43**, 1013–1020 (2005).
23. Groh, V., Wu, J., Yee, C. & Spies, T. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature* **419**, 734–738 (2002).
24. Doubrovina, E.S. *et al.* Evasion from NK cell immunity by MHC class I chain-related molecules expressing colon adenocarcinoma. *J. Immunol.* **171**, 6891–6899 (2003).
25. Kohga, K. *et al.* Serum levels of soluble major histocompatibility complex (MHC) class I-related chain A in patients with chronic liver diseases and changes during transcatheter arterial embolization for hepatocellular carcinoma. *Cancer Sci.* **99**, 1643–1649 (2008).
26. Jinushi, M. *et al.* Expression and role of MICA and MICB in human hepatocellular carcinomas and their regulation by retinoic acid. *Int. J. Cancer* **104**, 354–361 (2003).
27. Groh, V. *et al.* Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. *Proc. Natl. Acad. Sci. USA* **96**, 6879–6884 (1999).
28. Ota, M. *et al.* Trinucleotide repeat polymorphism within exon 5 of the MICA gene (MHC class I chain-related gene A): allele frequency data in the nine population groups Japanese, Northern Han, Hui, Uygur, Kazakhstan, Iranian, Saudi Arabian, Greek and Italian. *Tissue Antigens* **49**, 448–454 (1997).
29. Groh, V. *et al.* Costimulation of CD8αβ T cells by NKG2D via engagement by MIC induced on virus-infected cells. *Nat. Immunol.* **2**, 255–260 (2001).
30. Groh, V. *et al.* Cell stress-regulated human major histocompatibility complex class I gene expressed in gastrointestinal epithelium. *Proc. Natl. Acad. Sci. USA* **93**, 12445–12450 (1996).

ONLINE METHODS

Sample collections. We obtained DNA from 721 HCV-related HCC cases, 1,730 CHC cases and 5,486 HCV-negative controls from the BioBank Japan project[31]. For replication analysis, DNA from 673 HCV-induced HCC cases was obtained from a prospective HCC study cohort of the University of Tokyo. A diagnosis of CHC, liver cirrhosis or HCC were based on histological, clinical and laboratory findings obtained by trained physicians. Case samples with HBV co-infection were excluded from the analysis. Interferon was administrated to 20.4% of HCC cases and 70.1% of cases were not treated. The remaining 9.5% of the cases lacked information about interferon treatment. The non-HCV controls obtained from BioBank Japan contained case-mixed individuals after excluding all individuals with cancer, chronic hepatitis B, diabetes or tuberculosis. All subjects were of Japanese origin and provided written informed consent. The clinical and demographic details of the samples are summarized in **Supplementary Table 1**. We also obtained serum samples from BioBank Japan and the University of Tokyo (**Supplementary Table 12**). This research project was approved by the ethical committees of the University of Tokyo and RIKEN.

SNP genotyping and quality control. In the GWAS, 721 individuals with HCV-related liver cancer and 2,890 controls were genotyped using Illumina HumanHap610-Quad and Illumina HumanHap550v3 Genotyping BeadChip, respectively. In the replication stage, 673 cases with HCV-related disease, 1,730 cases with CHC and 2,596 controls were genotyped by the multiplex PCR-based Invader assay (Third Wave Technologies) and the Illumina HumanHap610-Quad, respectively. The common SNPs between the Illumina HumanHap550v3 and the Illumina HumanHap610-Quad arrays from all autosomal chromosomes were included for the analysis. We applied standard SNP quality control filters to exclude SNPs with low call rate (<99%), a Hardy-Weinberg equilibrium $P < 1.0 \times 10^{-6}$ for controls and minor allele frequency of <0.01. In the end, we obtained 432,703 SNPs for the analysis. In the replication analysis, the allele discrimination plots were validated by two well-trained researchers (the plots are available on request). We excluded samples with low genotyping rate (<99%) and employed principal component analysis to avoid the population stratification issue, in which individuals belonging only to Hondo cluster were included in the analysis (**Supplementary Fig. 10**)[32].

Statistical analysis. The association of SNPs with the disease phenotype in the GWAS, replication stage and combination analyses was tested using multivariate logistic regression analysis after adjusting for age at recruitment (continuous) and gender by assuming an additive model and using PLINK[33]. In the GWAS, the genetic inflation factor (λ) was derived by applying logistic regressed P values for all the tested SNPs. The quantile-quantile plot was drawn using R. The ORs were calculated by considering the major allele as a reference, unless it was stated otherwise elsewhere. The combined analysis of the GWAS and replication stage was verified by conducting the Mantel-Haenszel method. We considered $P < 5 \times 10^{-8}$ as the genome-wide significance threshold, which is the Bonferroni-corrected threshold for the number of independent SNPs genotyped in HapMap Phase 2 (ref. 34). Heterogeneity across the two stages was examined by using the Breslow-Day test[35].

For multiple logistic regression analysis at rs2596542 using the R program, we considered age at recruitment (≤60 or >60 years)[3], gender (male or female) and alcohol consumption (non-drinkers, ≤ 50 g alcohol per day or >50 g alcohol per day) as covariates from both the GWAS and replication stage cases with HCC and non-HCV controls. Association at the *MICA* VNTR locus was analyzed by Fisher's exact test, and the global P value was calculated using a χ^2 test. Statistical comparisons between genotypes and sMICA levels were performed by Kruskal-Wallis test or Wilcoxon rank test using R. We employed the R package haplo.stats to infer haplotypes and to perform haplotype association analysis. P values for association between sMICA levels and haplotype distribution were obtained by score test under an additive model by using the haplo.score function. ORs and 95% confidence intervals were calculated from the coefficients of the GLM model by considering the major haplotype as a reference. We used the haplo.cc function to calculate these statistical values.

HCV serotype. HCV serotype data was available for 531 cases with HCC from the replication stage. HCV serotype was examined by serotyping assay (SRL Laboratory) according to previously reported methods[36]. According to the Simmonds classification[37], serotype 1 corresponded to disease types 1a and 1b, whereas serotype 2 corresponded to disease types 2a and 2b.

***MICA* VNTR locus genotyping.** We followed the method suggested by Applied Biosystems. Briefly, the 5′ end of the forward primer was labeled with 6-FAM, and the 5′ end of reverse primer was labeled with the GTGTCTT non-random sequence to promote addition of As. The primer sequences were previously reported[28]. The PCR products were mixed with Hi-Di Formamide and GeneScan-600 LIZ size standard and separated using a GeneScan system on a 3730xl DNA analyzer (Applied Biosystems). GeneMapper software (Applied Biosystems) was used to assign the repeat fragment size (**Supplementary Fig. 7**).

Quantification of soluble MICA. sMICA levels were measured by sandwich enzyme-linked immunosorbent assay, as described in the manufacturer's instructions (R&D Systems).

Imputation and association analysis at HLA allele tagging SNPs. We obtained a SNP or a combination of SNPs which can tag HLA alleles in the Japanese population from a previous study[13]. The untyped genotypes of these SNPs were imputed in the GWAS samples by using a hidden Markov model programmed in MACH[38] and haplotype information from HapMap JPT samples. We applied the same SNP quality criteria as in the GWAS for selecting SNPs for the analysis. The association was tested on all SNPs that passed the quality control criteria using logistic regression analysis conditioned on age and gender.

Initially, we obtained the pair-wise LD between HLA alleles tagging SNPs and rs2596542. We performed case-control association analysis in our GWAS dataset. As shown in **Supplementary Table 9**, none of the HLA-tagging SNPs showed evidence of linkage or association except rs2844521, and rs2844521 was in absolute linkage with rs2596542 ($r^2 = 1$, $D' = 1$) and thus showed similar association. We obtained actual genotype data at rs2596501, as this SNP is included on the 550K SNP platform, and inferred the haplotype between rs2844521 and rs2596501. However, the haplotype GT (the G allele of rs2844521 and the T allele of rs2596501), which is reported to tag the *HLA-B*3501* allele ($r^2 = 1$, $D' = 1$), was not associated with HCC in our GWAS dataset ($P = 0.39$). We also performed a conditional logistic regression analysis on rs2596501 (data not shown) and found no effect on the association between rs2596542 and HCV-induced HCC. This data suggested that rs2596542 association is independent of *HLA-B*3501*. Although we observed mild association between other *HLA-B* alleles (*HLA-B*5401*, $P = 0.004$; *HLA-B*6701*, $P = 0.012$) and HCV-induced HCC, the association at rs2596542 alone was the most significant. Taken together, we found no strong evidence for linkage of HLA alleles with rs2596542.

Software. For general statistical analysis, we used R statistical environment version 2.6.1 or plink version 1.06. The Haploview software version 4.2 (ref. 39) was used to calculate LD and to draw Manhattan plots. Primer3 v0.3.0 web tool was used to design primers. We used LocusZoom for plotting regional association plots. We used FastSNP[40] web tool for functional annotation of SNPs (see URLs for all software packages).

31. Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).
32. Yamaguchi-Kabata, Y. *et al.* Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
33. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
34. Frazer, K.A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
35. Breslow, N. & Day, N. Statistical methods in cancer research. Volume II–The design and analysis of cohort studies. *IARC Sci. Publ.* 1–406 (1987).
36. Tsukiyama-Kohara, K. *et al.* A second group of hepatitis C viruses. *Virus Genes* **5**, 243–254 (1991).
37. Simmonds, P. *et al.* Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *J. Gen. Virol.* **75**, 1053–1061 (1994).
38. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
39. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
40. Yuan, H.Y. *et al.* FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.* **34**, W635–W641 (2006).

Hepatology Research 2010; **40**: 69–82 doi: 10.1111/j.1872-034X.2009.00593.x

Review Article

Animal models for hepatitis C and related liver disease

Kazuhiko Koike,[1] Kyoji Moriya[2] and Yoshiharu Matsuura[3]

Departments of [1]Gastroenterology and [2]Infection Control and Prevention, Graduate School of Medicine, University of Tokyo, Tokyo, and [3]Department of Molecular Virology, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan

Persistent infection with hepatitis C virus (HCV) is a major risk toward development of hepatocellular carcinoma (HCC). The elucidation of pathogenesis of HCV-associated liver disease is hampered by the absence of appropriate animal models: there has been no animal model for HCV infection/pathogenesis except for the chimpanzee. In contrast, a number of transgenic mouse lines carrying the cDNA of the HCV genome have been established and evaluated in the study of HCV pathogenesis. The studies using transgenic mouse models, in which the HCV proteins such as the core protein are expressed, indicate the direct pathogenicity of HCV, including oncogenic activities. HCV transgenic mouse models also show a close relationship between HCV and some hepatic and extrahepatic manifestations such as hepatic steatosis, insulin resistance or Sjögren's syndrome. A crucial role of hepatic steatosis and insulin resistance in the pathogenesis of liver disease in HCV infection has been demonstrated, implying hepatitis C to be a metabolic disease. Besides the data connecting liver fibrosis progression and the disturbance in lipid and glucose metabolisms in hepatitis C patients, a series of evidence was found showing the association between these two conditions and HCV infection, chiefly using transgenic mouse carrying the HCV genome. Furthermore, the persistent activation of peroxisome proliferator-activated receptor (PPAR)-α has recently been found, yielding dramatic changes in the lipid metabolism and oxidative stress overproduction in cooperation with the mitochondrial dysfunction. These results would provide a clue for further understanding of the role of lipid metabolism in pathogenesis of hepatitis C including liver injury and hepatocarcinogenesis.

Key words: core protein, hepatitis C, hepatocellular carcinoma, insulin resistance, steatosis, transgenic mouse.

INTRODUCTION

HEPATITIS C VIRUS (HCV) infection frequently evolves into a persistent state, leading to the development of chronic hepatitis, cirrhosis and, eventually, hepatocellular carcinoma (HCC). For understanding of the mechanism of entry into hepatocytes, replication and the pathogenesis of HCV, an *in vitro* replication system or animal models for HCV infection/pathogenesis have been eagerly awaited. An *in vitro* HCV replication system was not established until the development of a subgenomic, non-structural region HCV replicon system or an infectious genotype 2a HCV clone, JFH-1.[1] There has been no animal model for HCV infection/pathogenesis except for the chimpanzee.[2] Recently, however, several small animal models for HCV infection have been evaluated, including *Tupaia*[3] and genetically engineered mice that are chimeric for human hepatocytes.[4] On the other hand, a number of transgenic mouse lines carrying the cDNA of HCV genome have been established and evaluated in the study of HCV pathogenesis, as described hereafter. These mice, including those that are transgenic for the core gene of HCV, show the features resembling those of chronic hepatitis C patients, such as hepatic steatosis, insulin resistance and HCC. These animal models provide us a molecular understanding of the pathogenesis of HCV infection and a perspective for the future development of treatment and prophylaxis of liver disease occurring in HCV infection.

Correspondence: Professor Kazuhiko Koike, Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. Email: kkoike-tky@umin.ac.jp
Received 19 June 2009; revision 31 July 2009; accepted 31 July 2009.

THE CHIMPANZEE MODEL

AS EARLY AS the discovery of the cDNA clone of HCV, or even before that, the chimpanzee has been known to be susceptible to HCV (or the non-A, non-B

hepatitis agent), and has long been used as a sole animal model for HCV infection.[2] However, due to ethical reasons and vast costs, the use of this animal for HCV research is limited: the data on this animal model were obtained from the studies chiefly conducted in the USA. The serum samples from hepatitis C patients were inoculated to chimpanzees, and the natural course was evaluated in biochemical, virological or histological methods. These studies demonstrated that the course of HCV infection in this animal is similar to that in human beings, warranting the chimpanzee to be a good animal model for HCV infection, albeit HCC being a rare occurrence in chimpanzees.

In 1997, potential infectious HCV clones, which were produced by several study groups, were evaluated for *in vivo* infectivity using chimpanzees. The chimpanzees were also used for the evaluation of a role of cellular immunity in acute HCV infection: intrahepatic $CD4^+$ or $CD8^+$ T-cell response was found to play a crucial role in the eradication of HCV from the liver. Recently, this animal is also used for the evaluation of candidates for HCV vaccines and the assessment of *in vivo* infectivity of JFH-1 HCV viral clone, which shows a robust replication in human HCC-derived HuH-7 cells.[1] Immunization with virus-like particles of chimpanzees induced an HCV-specific immune response of $CD4^+$ or $CD8^+$ T cells, thereby suppressing the development of high viral loads in chimpanzees that were challenged with HCV.[5] Also, inoculation of the non-structural proteins of HCV using recombinant adenovirus vector induced HCV-specific immune T-cell response, leading to a significant suppression of replication of genotype 1a HCV that was challenged after immunization.[6]

In general, the liver lesions observed in HCV-inoculated chimpanzees are milder than those in human chronic hepatitis C patients, for example, cirrhosis or HCC rarely develops, but the morphological changes and inflammatory responses are similar to those in humans.[2] Therefore, the studies using chimpanzees are indispensable now and in the future for the analyses of viral replication, pathogenesis of liver disease and the evaluation of candidates for HCV vaccines.

THE SMALL PRIMATES MODEL

TUPAIA (*TUPAIA BELANGERI chinensis*), a small primate resembling the squirrel, has been reported to be susceptible to hepatitis B virus (HBV) infection in 1996,[3] and was used for the study of HCV infection.[7] However, only a quarter of inoculated individuals contracted HCV infection, and developed only a transient or intermittent viremia with low viral loads. Another study group reported on the usefulness of how a primary culture of hepatocytes from the liver of Tupaia can be infected with serum- or plasma-derived HCV from infected humans, as measured by de novo synthesis of HCV RNA, analysis of viral quasispecies evolution, and detection of viral proteins.[8,9]

While the development of liver disease (a cirrhosis-like lesion) in HCV-infected Tupaia was presented at scientific meetings, a scientific paper describing it has not appeared yet. In conclusion, the value of Tupaia in HCV research is limited, but it may be utilized for the analysis of viral entry or replication when HCV particles other than JFH-1 are used for the study.

HCV

THE DEVELOPMENT OF transgenic mouse technology was a great step forward in biotechnology in that this technology provides opportunities to examine *in vivo* an exceptionally wide variety of biological questions that were previously examined only *in vitro*. The selective addition of defined genes to the genome of a living animal is useful for investigating the consequences of expression of dominant genes, and thus a number of exogenous genes including oncogenes and humoral factor genes have been introduced into mouse eggs. Viral genes have also been transferred to define the complex cascades of events that can be triggered *in vivo* in response to the expression of a viral protein.

Hepatitis C virus is an enveloped RNA virus of the *Flavivirus* family, in which a positive-sense, single-stranded RNA genome of approximately 9600 nucleotides (nt) is contained within the nucleocapsid.[10] The genome consists of a large translational open reading frame (ORF) encoding a polyprotein of approximately 3010 amino acids (aa) (Fig. 1). The ORF is flanked by highly conserved untranslated regions (UTR) at both the 5′ and 3′ termini. The complete 5′ UTR consists of 341 nt and acts as an internal ribosomal entry site. This feature leads to the translation of the RNA genome using a cap-independent mechanism, rather than ribosome scanning from the 5′ end of a capped molecule.

The polyprotein is processed by both the cellular and viral proteases to generate the viral gene products, which are subdivided into the structural and non-structural proteins. The structural proteins, which are encoded by the NH_2-terminal quarter of the genome, include the

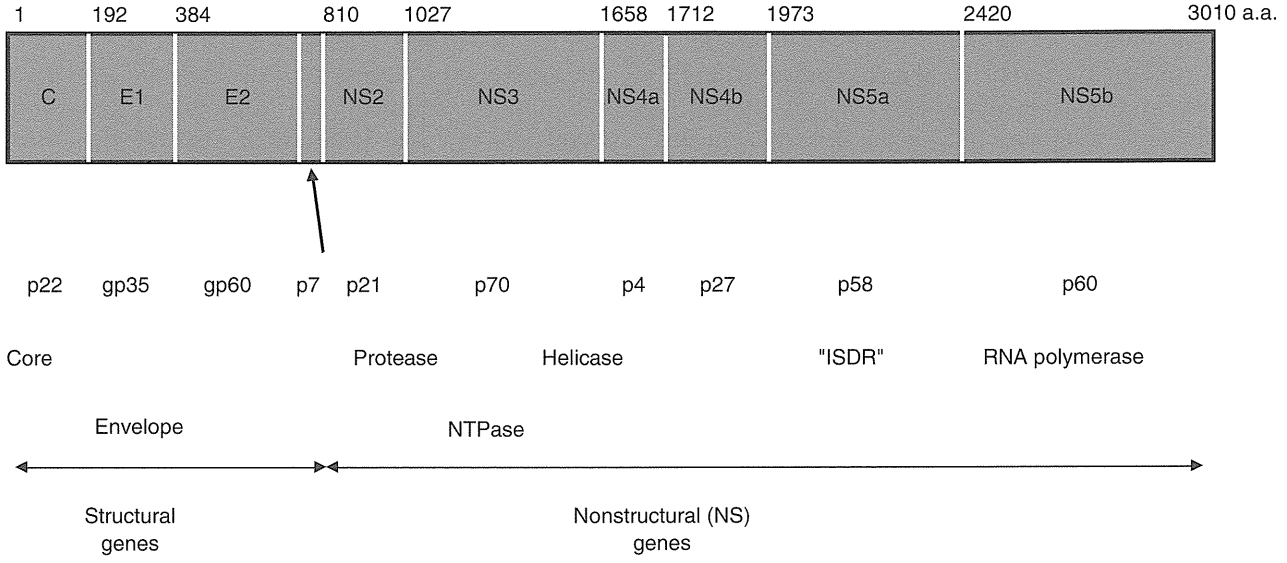


Figure 1 The structure of hepatitis C virus (HCV) genome. The HCV genome RNA encodes a polyprotein of 3010 amino acids (a.a.), which is processed to structural and non-structural proteins by the cellular or viral proteases. ISDR, interferon sensitivity-determining region.

core protein and the envelope proteins, E1 and E2. The E2 has an alternative form, E2-p7, though it is not clear whether or not the p7 composes the viral particle. The NS2, NS3, NS4A, NS4B, NS5A and NS5B are the non-structural proteins that are coded in the remaining portion of the polyprotein. These include serine protease (NS3/4A), NTPase/helicase (NS3) and RNA-dependent RNA polymerase (NS5B).

The core protein of HCV occupies residues 1–191 of the precursor polyprotein and is cleaved between the core and E1 protein by host signal peptidase. The C-terminal membrane anchor of the core protein is further processed by host signal peptide peptidase.[11] The mature core protein is estimated to consist of 177–179 amino acids and shares high homology among HCV genotypes. The HCV core protein possesses the hydrophilic N-terminal region "domain 1" (residues 1–117) followed by a hydrophobic region called "domain 2", which is located from residue 118–170. The domain 1 is rich in basic residues, and is implicated in RNA-binding and homo-oligomerization. The amphipathic helices I and II spanning from residue 119–136 and residue 148–164, respectively, in domain 2 are involved in the association of HCV core protein with lipid.[12] In addition, the region spanning from residue 112–152 is associated with membranes of the endoplasmic reticulum and mitochondria.[13] The core protein is also localized into the nucleus[14,15] and binds to the nuclear proteasome activator PA28γ/REGγ, resulting in PA28γ-dependent degradation of the core protein.[16]

A recent report suggests that ubiquitination and adenosine triphosphate (ATP) are not required for PA28γ-dependent proteasome activity.[17] HCV core protein is also known to be ubiquitinated by E3 ligase E6AP and degraded in the ubiquitin/ATP-dependent pathway.[14] Thus, the HCV core protein is degraded in at least two different ways. To further assess the pathological significance of the interaction of core protein with PA28γ, Core-Tg/PA28γ-knockout mice have been generated and analyzed as described below (section 9).[15]

POSSIBLE ROLE OF HCV IN HEPATOCARCINOGENESIS

THE MECHANISM UNDERLYING hepatocarcinogenesis in HCV infection is not fully understood yet, despite the fact that nearly 80% of patients with HCC in Japan are persistently infected with HCV.[18–20] HCV infection is also common in patients with HCC in other countries albeit to a lesser extent. These lines of evidence prompted us to seek to determine the role of HCV in hepatocarcinogenesis. Inflammation induced by HCV should be considered, of course, in a study on the hepatocarcinogenesis in hepatitis viral infection: necrosis of hepatocytes due to chronic inflammation followed by regeneration enhances genetic aberrations in host cells,

the accumulation of which culminates in HCC. This theory presupposes an indirect involvement of hepatitis viruses in HCC through hepatic inflammation. However, this context leaves us with a serious question: can inflammation alone result in the development of HCC in such a high incidence (90% in 15 years) or the multicentric nature of HCV infection?

The other role of HCV would have to be weighed against a rare occurrence of HCC in patients with autoimmune hepatitis in which severe inflammation in the liver persists indefinitely, even after the development of cirrhosis. These backgrounds and reasonings lead to a possible activity of viral proteins for inducing neoplasia. This possibility has been evaluated by introducing genes of HCV into hepatocytes in culture with little success. One of the difficulties in using cultured cells is the carcinogenic capacity of HCV, if any, which would be weak and would take a long time to manifest itself. Actually, it takes 30–40 years for HCC to develop in individuals infected with HCV. On the basis of these viewpoints, we started to investigate carcinogenesis in chronic hepatitis C, *in vivo*, by transgenic mouse technology.

TRANSGENIC MOUSE LINES CARRYING THE HCV GENOME

AS DESCRIBED ABOVE, the HCV proteins have been characterized chiefly using *in vitro* translation or cultured cells. Little is known, however, about the role of HCV or its proteins in the pathogenesis of hepatitis and subsequent liver diseases, cirrhosis and HCC. One of the major issues regarding the pathogenesis of HCV-associated liver lesions is whether the HCV proteins have direct effects on pathological phenotypes. Although several strategies have been used to characterize the hepatitis C viral proteins, the relationship between the protein expression and disease phenotype has not been clarified. For this purpose, several lines of mice have been established which were transgenic for the HCV cDNA (Table 1). They include the ones carrying the entire coding region of the HCV genome,[33] the core region only,[21,29] the envelope region only,[30,31] the core and envelope regions[33,34] and the core to NS2 regions.[41] Although detection of mRNA from the NS regions of the HCV cDNA has been reported,[33,37] the detection of HCV NS proteins in the transgenic mouse liver have not been successful. The reason for this failure in detecting NS proteins is unclear, but the expression of the NS enzymes may be harmful to mouse development and may allow the establishment of only low-expression mice.

In terms of expression system, two different ways have been applied; transient and constitutive expression systems. One transgenic mouse line has been reported which expresses the HCV genes using a transient expression system. Wakita *et al.* utilized the Cre/*lox*P system, by which a gene under silent can be switched on by the introduction of Cre recombinase. They established a transgenic mouse line that had the core, envelopes and NS2 genes of HCV in a silent state. After the injection of the recombinant adenovirus that had *Cre* recombinase in the mice, the HCV genes expressed transiently.[41] These mice developed acute hepatitis, which was blocked by the administration of anti-CD4 and CD8 antibodies. This mouse system would provide a good animal model for acute hepatitis C and be useful for the study of immunological aspects of hepatitis. The possibility, however, that the greatly overexpressed HCV proteins had caused the death of hepatocytes and provoked the immune response thereafter still remains.

We have engineered transgenic mouse lines carrying the HCV genome by introducing the genes from the cDNA of the HCV genome of genotype 1b.[21,22] Established are three different kinds of transgenic mouse lines, which carry the core gene, envelope genes or non-structural genes, respectively, under the same transcriptional regulatory element. Among these mouse lines, only the transgenic mice carrying the core gene developed HCC in two independent lineages (Fig. 2).[22] The envelope gene transgenic mice do not develop HCC, despite high expression levels of both E1 and E2 proteins,[31,32] and the transgenic mice carrying the entire non-structural genes have developed no HCC.

The core gene transgenic mice express the core protein of an expected size, and the level of the core protein in the liver is similar to that in chronic hepatitis C patients. Early in life, these mice develop hepatic steatosis, which is one of the histological characteristics of chronic hepatitis C, along with lymphoid follicle formation and bile duct damage.[43] Thus, the core gene transgenic mouse model well reproduces the feature of chronic hepatitis C. Of note, any pictures of significant inflammation are not observed in the liver of this animal model. Late in life, these transgenic mice develop HCC. Notably, the development of steatosis and HCC has been reproduced by other HCV transgenic mouse lines, which harbor the entire HCV genome or structural genes including the core gene.[29,33,34] These outcomes indicate that the core protein per se of HCV has an oncogenic potential when expressed *in vivo*.