Table 3 Demographic, virological and clinical features of patients with chronic hepatitis C treated by IFN plus ribavirin combination therapy

Variable	SVR (n = 98)*	Non-SVR (n = 105)*	P-value
Sex			
Male	64	50	**0.011**
Female	34	55	
Age (years)†	56.0 (24–72)	58.5 (27–74)	**0.023**
Weight (kg)†	61.8 (43–91)	61.9 (41–94)	0.821
Pre-treatment ALT (IU/L)†	69 (17–285)	57 (16–304)	0.170
Interferon history‡			
Naive	58 (51.8%)	54 (48.2%)	0.311
Relapse	30 (47.6%)	33 (52.4%)	
Nonresponse	10 (35.7%)	18 (64.3%)	
HCV genotype			
1	50 (34.0%)	97 (66.0%)	**0.00000002**
2, 3	48 (85.7%)	8 (14.3%)	
HCV RNA titre (kIU/mL)			
<100	15 (93.7%)	1 (6.3%)	**0.004**
100–500	35 (49.3%)	36 (50.7%)	
500–850	22 (44.9%)	27 (55.1%)	
850≤	26 (38.8%)	41 (61.2%)	
Fibrosis score			
0	5 (62.5%)	3 (37.5%)	**0.0005**
1	38 (59.4%)	26 (40.6%)	
2	19 (45.2%)	23 (54.8%)	
3	5 (20.0%)	20 (80.0%)	
4	2 (28.6%)	5 (71.6%)	

SVR, sustained virologic response. P-values in boldface are significant.

*SVR and non-SVR were evaluated in patients who had completed therapy for 24 or 48 weeks.

†Values are median (range).

‡One hundred six patients had received previous treatment with IFN-α monotherapy for 24 weeks, but failed to respond or relapsed.

HCV genotypes 2 and 3, 48 (85.7%) had SVR, whereas 50 of 147 (34.0%) patients with HCV genotype 1 had SVR, indicating that HCV genotype 1 was significantly associated with non-SVR (P = 0.00000002). In addition, a lower viral load before treatment (P = 0.004), male sex (P = 0.011), young age (P = 0.023), and lower degree of liver fibrosis (P = 0.0005) were significantly associated with SVR.

SNP genotyping analyses revealed that the frequencies of all 35 polymorphisms detected in the 240 hepatitis C patients were not significantly different from those in healthy volunteers. The success scores of the Taqman assay were 96.4–100% and those of direct sequencing were 95.8–100%. Univariate analyses of 35 polymorphisms revealed that a *TYK2* exon8 15560-G/T polymorphism (rs2304256) was significantly associated with virologic response to IFN-based therapy [P = 0.050, OR = 0.66 (0.44–0.99)] (Table S5).

In contrast to the univariate analysis, however, multiple logistic regression analyses demonstrated that the rs2304256 was not significant (P = 0.675) (Table 4). As a host factor, only a lower degree of liver fibrosis before therapy (P = 0.007) was significantly associated with SVR in the multiple logistic regression model. On the other hand, HCV genotypes 2 and 3 (P = 0.00005) and a lower viral load before therapy (P = 0.027) were both significantly associated with SVR.

Genetic polymorphisms associated with the adverse effects of IFN-based therapy

A total of 132 of 240 (55.0%) patients required either a discontinuation or a dose reduction of IFN or RBV due to the following adverse events: anaemia (n = 50), neutropenia or leucocytopenia (n = 32), thrombocytopenia (n = 17), depression (n = 7), and other causes (malaise, alopecia, and abdominal discomfort). The relationship between baseline characteristics and occurrence of haematologic adverse effects of the IFN plus RBV combination therapy is summarized in Table S6.

To identify the host genetic polymorphisms associated with the haematologic adverse effects of IFN plus RBV therapy, we focused on decreases in blood cell counts during the therapy and analysed the association with the SNPs in IFN signalling pathway-related genes. Consistent with previous reports [30,31], leucocyte, neutrophil, and platelet counts and haemoglobin levels usually declined in the initial 2–4 weeks of treatment, then stabilized during treatment, and returned to baseline levels within 12 weeks from the end of treatment in patients receiving IFN plus RBV therapy (Fig. 1). Therefore, we evaluated the decreases in leucocyte, neutrophil, and platelet counts and haemoglobin level at 4 weeks of treatment. We first examined the predictive factors for neutropenia. In 240 patients, absolute neutrophil counts decreased by an average of 39.3% from baseline during the first 4 weeks of treatment. Univariate analyses of 32 polymorphisms and clinical features showed that two SNPs, an *IFNAR1* intron2 10848-A/G polymorphism (rs2243594), and a *STAT2* intron5 4757-G/T, were associated with neutropenia caused by IFN-based therapy [P = 0.038, P = 0.020] (Table 5, Table S7). Furthermore, multivariate linear regression analysis confirmed that both polymorphisms were significantly associated with the neutropenia (P = 0.013, P = 0.009). Next, we examined the predictive factors for leucocytopenia. Absolute leucocyte counts decreased by an average of 29.9% from baseline within the first 4 weeks of treatment. Univariate analyses indicated that an *IFNAR1* intron2 10848-A/G polymorphism (rs2243594), an *IRF2* intron6 66675-C/T polymorphism (rs2241500), and female sex were associated with leucocytopenia (P = 0.048, P = 0.026, P = 0.016,

Table 4 Univariate and multiple logistic regression analyses of SNPs and clinical factors associated with the efficacy of IFN plus ribavirin combination therapy

	Univariate analysis		Multiple logistic regression analysis	
Variable	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)
SNPs				
TYK2 15660-G/T	**0.050**	0.66 (0.44–0.99)	0.675	0.48 (0.14–1.67)
Clinical variables				
Sex	**0.011**	2.07 (1.17–3.66)	0.082	2.09 (0.90–4.84)
Age	**0.023**	0.16 (0.04–0.65)	0.347	0.69 (0.11–4.22)
HCV genotype	**0.00000002**	11.6 (5.09–26.6)	**0.00005**	7.35 (2.54–21.2)
Viral load	**0.004**	0.25 (0.10–0.62)	**0.027**	0.22 (0.06–0.88)
Fibrosis stage	**0.0005**	12.0 (2.63–54.8)	**0.007**	10.3 (1.72–62.3)

P-values in boldface are significant. SNP, single nucleotide polymorphism.

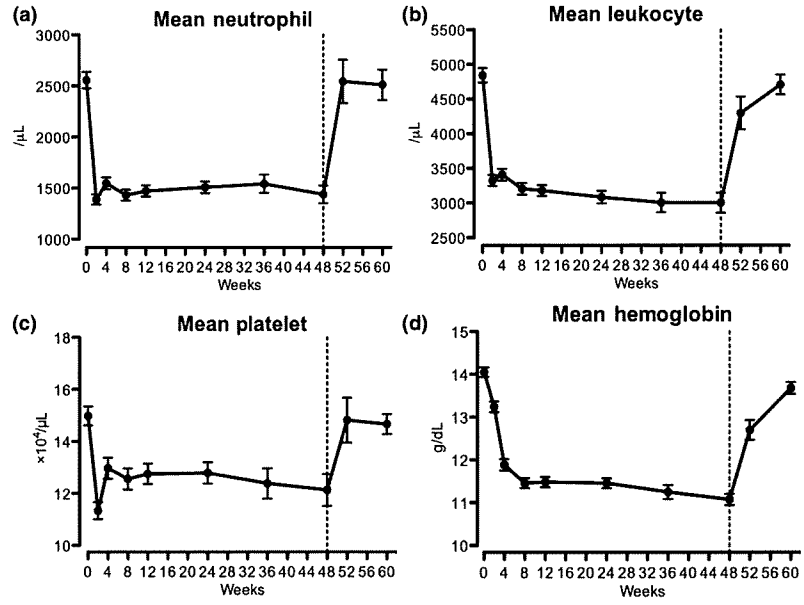

Fig. 1 Change in mean neutrophil (a), leucocyte (b), and platelet counts (c), and haemoglobin levels (d) during and after IFN plus RBV therapy. The results are shown as mean ± SEM.

respectively). Multivariate analysis, however, indicated that none of the factors, including *IFNAR1* rs2243594, *IRF2* rs2241500 and sex, were significant. Third, we examined the predictive factors for thrombocytopenia. Absolute platelet counts decreased by an average of 12.5% from baseline during the first 4 weeks of treatment. Univariate analyses showed that only an *IRF7* exon2 789-G/A (rs1061501) was associated with thrombocytopenia ($P = 0.031$). Finally, we examined the predictive factors for anaemia. Absolute haemoglobin concentration decreased by an average of 15.8% of baseline within the first 4 weeks of treatment. Univariate analyses revealed that anaemia was associated with older age ($P = 0.0004$), but not with any of the polymorphisms.

We examined the genotype results (variant allele carrier) of an *IFNAR1* intron2 10848-A/G polymorphism (rs2243594), a *STAT2* intron5 4757-G/T polymorphism, and an *IRF7* exon2 789-G/A polymorphism (rs1061501) for their association with various clinical and histologic features among 240 patients (Table S8). None of the factors, however, were associated with the SNPs identified.

DISCUSSION

In this study, we evaluated the influence of genetic polymorphisms on adverse effects and efficacy of IFN plus RBV combination therapy. Although several studies have evaluated the influence of host genetic polymorphisms on virologic response to IFN-based therapy, no studies have looked at possible association of adverse effects of the IFN-based therapy and host genetic polymorphisms. We report for the first time that certain SNPs in the IFN signalling pathway-related genes were associated with haematologic adverse effects in chronic hepatitis C patients undergoing IFN-based therapy.

Table 5 Univariate and multiple linear regression analyses of SNPs and clinical factors associated with leucocytopenia, neutropenia and thrombocytopenia

Variable	Unit of B coefficient	Univariate analysis			Multiple analysis		
		P-value	*B* coefficient	SE *B*	*P*-value	*B* coefficient	SE *B*
Neutropenia							
IFNAR1 10848-A/G	%	**0.038**	6.94	3.31	**0.013**	6.43	2.57
STAT2 4757-G/T	%	**0.020**	−14.3	6.09	**0.011**	−13.8	5.41
Leucocytopenia							
IFNAR1 10848-A/G	%	**0.048**	4.14	2.08	0.109	1.62	1.61
IRF2 66675-C/T	%	**0.026**	3.44	1.53	0.054	3.00	1.54
Sex	%	**0.016**	7.79	3.20	0.134	3.30	2.20
Thrombocytopenia							
IRF7 789-G/A	%	**0.031**	4.15	1.92	ND	ND	ND
Anaemia							
Age	%/year	**0.0004**	0.28	0.08	ND	ND	ND

P-values in boldface are significant. SE, standard error. ND, not done because only one factor was significant in the univariate analysis.

The representative side effect of IFN-based combination therapy with RBV that causes poor therapeutic tolerance is haematologic toxicity, such as anaemia, neutropenia, and thrombocytopenia [4,32]. In fact, several studies reported that less than half the patients with hepatitis C were able to complete IFN plus RBV combination therapy at the assigned dose of both drugs, causing reduced therapeutic efficacy [5,6]. One thing to be noted is that the decrease in neutrophil and platelet counts induced by IFN-based therapy varies among patients, and thus it is difficult to predict the risk of haematologic toxicities in chronic hepatitis C patients receiving IFN-based therapy. The molecular mechanism of IFN-induced haematologic toxicities, however, is unknown. Several studies suggested the possibility that IFN treatment causes bone marrow suppression [33,34]. In agreement with this hypothesis, it was shown that a significant drop in platelet count after the initiation of IFN therapy is accompanied by a moderate increase in thrombopoietin levels in the failing liver, which may be insufficient to counteract the myelosuppressive action of IFN [35]. Another study suggested that IFN-mediated cytopenia may be due to rapid sequestration of platelets and leucocytes in the capillary beds of the liver and spleen [36]. Our current findings suggest that some of the IFN signalling pathway-related genes are involved in the decrease in neutrophil and platelet counts in response to IFN treatment. Interestingly, a recent study demonstrated that an intrinsic program for apoptosis controls platelet survival and dictates life span [37]. They revealed that platelets are genetically programmed to die by apoptosis and the antagonistic balance between antiapoptotic and proapoptotic molecules determines platelet life span. It is well known that IFN signalling induces the expression of multiple IFN-stimulated genes including molecules with proapoptotic or antiapoptotic function, such as tumour necrosis factor-related apoptosis-inducing ligand Fas, and X-linked inhibitor of apoptosis-associated factor 1 [38]. Thus, it is possible that IFNAR1, STAT2, and IRF7 contribute to the occurrence of neutropenia and thrombocytopenia by regulating the magnitude of IFN signalling involved in the apoptotic pathway in the haematopoietic cells in patients receiving IFN-based treatment.

In this study, three SNPs were associated with cytopenia in chronic hepatitis C patients receiving IFN plus RBV combination therapy. Among them, rs1061501 in the *IRF7* gene was located in the exon region but is a synonymous SNP. Recently, Kimchi-Sarfaty *et al.* demonstrated [39] that a synonymous SNP that did not affect amino acid sequence was capable of changing the function of the resultant protein. Indeed, the presence of a rare codon marked by a synonymous SNP in the *Multidrug Resistance 1* gene affects the timing of cotranslational folding and thereby alters the structure of substrate. Thus, it is possible that the synonymous rs1061501 contributes to a functional change in the IRF7 protein. On the other hand, rs2243594 in the *IFNAR1* gene and the SNP in the *STAT2* gene associated with neutropenia were located in an intronic region. In general, intronic SNPs provide little evidence for changes in protein structure or function, but an intronic mutation in the *p53* gene could have functional consequences by regulating gene expression, suggesting that the effect is mediated by a nonsynonymous and disruptive coding change in linkage disequilibrium with the associated intronic SNP or by a change in RNA splicing, editing, or expression [40]. Thus, it is possible that two intronic SNPs associated with neutropenia contribute to functional changes in the IFNAR1 and STAT2 proteins.

In contrast to the adverse effects of IFN plus RBV combination therapy, none of the host genetic polymorphisms in the IFN signalling pathway-related genes analysed were

associated with therapeutic efficacy. The results indicated that viral factors, including viral genotype and pre-treatment viral load, and histological fibrosis grade were likely to have critical roles in treatment response. Consistent with many previous reports [41–43], we found that HCV genotypes 2 and 3, low viral load, and early fibrosis stage predict a favourable virologic response to IFN plus RBV combination therapy. On the other hand, it was reported that several SNPs in certain genes are associated with efficacy in IFN-based therapy [8–14,16,17]. Many of these previous studies, however, evaluated the association between the SNP and the treatment response using only univariate and not multivariate analyses that included viral factors. In fact, in our univariate analysis, one *TYK2* SNP (rs2304256) showed a possible association with therapeutic efficacy. Multivariate analysis, however, revealed that this SNP was not significant. Taken together, these findings suggest that the viral factors and host histological grade of liver fibrosis are important predictors of the treatment response in chronic hepatitis C infection. Although no significant association was observed between the efficacy and the IFN signalling pathway-related genes examined, it is possible that polymorphisms of other genes might play a role in the treatment response to IFN-based therapy.

In conclusion, we demonstrated that the SNPs in the *IFNAR1* and *STAT2* genes were associated with neutropenia and the SNP in the *IRF7* gene was associated with thrombocytopenia in chronic hepatitis C patients receiving IFN plus RBV combination therapy. In contrast, the virologic factors and histological grade of liver fibrosis are important predictors for virologic response to the IFN-based therapy, whereas no host genetic polymorphisms in IFN signalling pathway-related genes analysed affected the therapeutic efficacy. Further analyses are required to clarify the mechanisms of how those polymorphisms affect the biologic function of the IFN signalling and contribute to the occurrence of haematological adverse effects in IFN-treated patients.

ACKNOWLEDGEMENTS

The following institutions and investigators also participated in the study: Dr T. Nakamura, Tenri Hospital; Dr T. Inokuma, Kobe City Medical Center General Hospital; Dr K. Ikeda, Japan Baptist Hospital; Dr A. Nakamura, Sanda City Hospital; Dr H. Yamada, Shinko Hospital; Dr H. Komori, Nishi Kobe Medical Center; Dr T. Tamada, Takatsuki Red Cross Hospital; Dr Y. Yamashita, Miki City Hospital; Dr K. Mizuta, Shiga Medical Center for Adults and Dr N. Kitajima, Kasai City Hospital.

This work was supported by Grants-in-aid for Scientific Research 16017240, 16017249, 17013051, 17659212 and 18012029 from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Grant-in-aid for Scientific Research 15209024 and 18209027 from JSPS, and Grant-in-Aid for Research in Measures for Intractable Diseases, and Research in Advanced Medical Technology (nano0051) from the Ministry of Health, Labor, and Welfare, Japan.

REFERENCES

1 Lauer GM, Walker BD. Hepatitis C virus infection. *N Engl J Med* 2001; 345: 41–52.
2 National Institutes of Health Consensus Development Conference Statement. Management of hepatitis C: 2002 – June 10–12, 2002. *Hepatology* 2002; 36: S3–20.
3 Dienstag JL, McHutchison JG. American Gastroenterological Association technical review on the management of hepatitis C. *Gastroenterology* 2006; 130: 231–264; quiz 214–237.
4 Hoofnagle JH, Seeff LB. Peginterferon and ribavirin for chronic hepatitis C. *N Engl J Med* 2006; 355: 2444–2451.
5 Fried MW, Shiffman ML, Reddy KR *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002; 347: 975–982.
6 Manns MP, McHutchison JG, Gordon SC *et al.* Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 2001; 358: 958–965.
7 Hadziyannis SJ, Sette H Jr, Morgan TR *et al.* Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann Intern Med* 2004; 140: 346–355.
8 Matsuyama N, Mishiro S, Sugimoto M *et al.* The dinucleotide microsatellite polymorphism of the IFNAR1 gene promoter correlates with responsiveness of hepatitis C patients to interferon. *Hepatol Res* 2003; 25: 221–225.
9 Edwards-Smith CJ, Jonsson JR, Purdie DM, Bansal A, Shorthouse C, Powell EE. Interleukin-10 promoter polymorphism predicts initial response of chronic hepatitis C to interferon alfa. *Hepatology* 1999; 30: 526–530.
10 Yee LJ, Tang J, Gibson AW, Kimberly R, Van Leeuwen DJ, Kaslow RA. Interleukin 10 polymorphisms as predictors of sustained response in antiviral therapy for chronic hepatitis C infection. *Hepatology* 2001; 33: 708–712.
11 Dai CY, Chuang WL, Chang WY *et al.* Tumor necrosis factor-alpha promoter polymorphism at position −308 predicts response to combination therapy in hepatitis C virus infection. *J Infect Dis* 2006; 193: 98–101.
12 Huang Y, Yang H, Borg BB *et al.* A functional SNP of interferon-gamma gene is important for interferon-alpha-induced and spontaneous recovery from hepatitis C virus infection. *Proc Natl Acad Sci U S A* 2007; 104: 985–990.
13 Promrat K, McDermott DH, Gonzalez CM *et al.* Associations of chemokine system polymorphisms with clinical outcomes and treatment responses of chronic hepatitis C. *Gastroenterology* 2003; 124: 352–360.
14 Naito M, Matsui A, Inao M *et al.* SNPs in the promoter region of the osteopontin gene as a marker predicting the efficacy of interferon-based therapies in patients with chronic hepatitis C. *J Gastroenterol* 2005; 40: 381–388.
15 Schott E, Witt H, Neumann K *et al.* Association of TLR7 single nucleotide polymorphisms with chronic HCV-

infection and response to interferon-a-based therapy. *J Viral Hepatol* 2008; 15: 71–78.
16 Wasmuth HE, Werth A, Mueller T *et al.* Haplotype-tagging RANTES gene variants influence response to antiviral therapy in chronic hepatitis C. *Hepatology* 2004; 40: 327–334.
17 Bonkovsky HL, Naishadham D, Lambrecht RW *et al.* Roles of iron and HFE mutations on severity and response to therapy during retreatment of advanced chronic hepatitis C. *Gastroenterology* 2006; 131: 1440–1451.
18 McHutchison JG, Gordon SC, Schiff ER *et al.* Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med* 1998; 339: 1485–1492.
19 Borden EC, Sen GC, Uze G *et al.* Interferons at age 50: past, current and future impact on biomedicine. *Nat Rev Drug Discov* 2007; 6: 975–990.
20 Gao B, Hong F, Radaeva S. Host factors and failure of interferon-alpha treatment in hepatitis C virus. *Hepatology* 2004; 39: 880–890.
21 de Veer MJ, Holko M, Frevel M *et al.* Functional classification of interferon-stimulated genes identified using microarrays. *J Leukoc Biol* 2001; 69: 912–920.
22 Gale M Jr, Foy EM. Evasion of intracellular host defence by hepatitis C virus. *Nature* 2005; 436: 939–945.
23 Kawai T, Akira S. Innate immune recognition of viral infection. *Nat Immunol* 2006; 7: 131–137.
24 Honda K, Taniguchi T. IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. *Nat Rev Immunol* 2006; 6: 644–658.
25 Stellacci E, Testa U, Petrucci E *et al.* Interferon regulatory factor-2 drives megakaryocytic differentiation. *Biochem J* 2004; 377: 367–378.
26 Mizutani T, Tsuji K, Ebihara Y *et al.* Homeostatic erythropoiesis by the transcription factor IRF2 through attenuation of type I interferon signaling. *Exp Hematol* 2008; 36: 255–264.
27 The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology* 1994; 20: 15–20.
28 Takahashi M, Matsuda F, Margetic N, Lathrop M. Automated identification of single nucleotide polymorphisms from sequencing data. *J Bioinform Comput Biol* 2003; 1: 253–265.
29 Sebastiani P, Lazarus R, Weiss ST, Kunkel LM, Kohane IS, Ramoni MF. Minimal haplotype tagging. *Proc Natl Acad Sci U S A* 2003; 100: 9900–9905.
30 Fried MW. Side effects of therapy of hepatitis C and their management. *Hepatology* 2002; 36: S237–S244.
31 Matthews SJ, McCoy C. Peginterferon alfa-2a: a review of approved and investigational uses. *Clin Ther* 2004; 26: 991–1025.
32 Manns MP, Wedemeyer H, Cornberg M. Treating viral hepatitis C: efficacy, side effects, and complications. *Gut* 2006; 55: 1350–1359.
33 Ernstoff MS, Kirkwood JM. Changes in the bone marrow of cancer patients treated with recombinant interferon alpha-2. *Am J Med* 1984; 76: 593–596.
34 Ganser A, Carlo-Stella C, Greher J, Volkers B, Hoelzer D. Effect of recombinant interferons alpha and gamma on human bone marrow-derived megakaryocytic progenitor cells. *Blood* 1987; 70: 1173–1179.
35 Peck-Radosavljevic M, Wichlas M, Pidlich J *et al.* Blunted thrombopoietin response to interferon alfa-induced thrombocytopenia during treatment for hepatitis C. *Hepatology* 1998; 28: 1424–1429.
36 Dormann H, Krebs S, Muth-Selbach U *et al.* Rapid onset of hematotoxic effects after interferon alpha in hepatitis C. *J Hepatol* 2000; 32: 1041–1042.
37 Mason KD, Carpinelli MR, Fletcher JI *et al.* Programmed anuclear cell death delimits platelet life span. *Cell* 2007; 128: 1173–1186.
38 Chawla-Sarkar M, Lindner DJ, Liu YF *et al.* Apoptosis and interferons: role of interferon-stimulated genes as mediators of apoptosis. *Apoptosis* 2003; 8: 237–249.
39 Kimchi-Sarfaty C, Oh JM, Kim IW *et al.* A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* 2007; 315: 525–528.
40 Lehman TA, Haffty BG, Carbone CJ *et al.* Elevated frequency and functional activity of a specific germ-line p53 intron mutation in familial breast cancer. *Cancer Res* 2000; 60: 1062–1069.
41 Liang TJ, Rehermann B, Seeff LB, Hoofnagle JH. Pathogenesis, natural history, treatment, and prevention of hepatitis C. *Ann Intern Med* 2000; 132: 296–305.
42 Trepo C. Genotype and viral load as prognostic indicators in the treatment of hepatitis C. *J Viral Hepat* 2000; 7: 250–257.
43 Tsubota A, Chayama K, Ikeda K *et al.* Factors predictive of response to interferon-alpha therapy in hepatitis C virus infection. *Hepatology* 1994; 19: 1088–1094.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1 Oligonucleotide sequences for primers and probes used for Taqman SNP genotyping assay.

Table S2 Assay ID of primers and probes used for Taqman SNP genotyping assay.

Table S3 Oligonucleotide sequences for primers used for PCR amplification and sequencing.

Table S4 List of discovered polymorphisms in 12 IFN-signalling related genes.

Table S5 Genotype frequency in the genotyped 35 polymorphisms of the IFN-signalling related genes.

Table S6 Demographic, virological and clinical features of patients with chronic hepatitis C treated by IFN plus ribavirin combination therapy.

Table S7 Linear regression analyses of 35 SNPs and clinical factors associated with haematologic adverse effects.

Table S8 Demographic, and clinical features according to three polymorphisms significantly associated with IFN-induced neutropenia and thrombocytopenia.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Oncogene (2009) 28, 469–478

www.nature.com/onc

ORIGINAL ARTICLE

A novel mouse model of hepatocarcinogenesis triggered by AID causing deleterious p53 mutations

A Takai[1], T Toyoshima[3], M Uemura[3], Y Kitawaki[2], H Marusawa[1], H Hiai[3], S Yamada[4], IM Okazaki[2], T Honjo[2], T Chiba[1] and K Kinoshita[3]

[1]Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [2]Department of Immunology and Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [3]Shiga Medical Center Research Institute, Moriyama, Japan and [4]Mikasa Laboratory, Immuno-Biological Laboratories Co., Ltd, Mikasa, Japan

Activation-induced cytidine deaminase (AID), the only enzyme that is known to be able to induce mutations in the human genome, is required for somatic hypermutation and class-switch recombination in B lymphocytes. Recently, we showed that AID is implicated in the pathogenesis of human cancers including hepatitis C virus (HCV)-induced human hepatocellular carcinoma (HCC). In this study, we established a new AID transgenic mouse model (TNAP-AID) in which AID is expressed in cells producing tissue-nonspecific alkaline phosphatase (TNAP), which is a marker of primordial germ cells and immature stem cells, including ES cells. High expression of TNAP was found in the liver of the embryos and adults of TNAP-AID mice. HCC developed in 27% of these mice at the age of approximately 90 weeks. The HCC that developed in TNAP-AID mice expressed α-fetoprotein and had deleterious mutations in the tumour suppressor gene *Trp53*, some of which corresponded to those found in human cancer. In conclusion, TNAP-AID is a mouse model that spontaneously develops HCC, sharing genetic and phenotypic features with human HCC, which develops in the inflamed liver as a result of the accumulation of genetic changes.
Oncogene (2009) **28,** 469–478; doi:10.1038/onc.2008.415; published online 10 November 2008

Keywords: hepatic cancer; stem cell marker; activation-induced cytidine deaminase; Cre recombinase; somatic mutation; animal model

Introduction

It is widely recognized that mutations of oncogenes, tumour suppressor genes and genomic stability genes play pivotal roles in cancer development (Vogelstein and Kinzler, 2004). Despite remarkable progress in our knowledge of the molecular mechanisms of individual cancer-related genes, surprisingly little is known about the fundamental aspects of when and how mutations are introduced into what kind of cell populations (for example, differentiated cells versus tissue stem cells).

To address this problem, a new mechanism of mutagenesis for cancer development has recently been proposed (Kinoshita and Nonaka, 2006; Marusawa, 2008). It hypothesizes that at least some cancer-related mutations are introduced by activation-induced cytidine deaminase (AID), an enzyme that is expressed in activated B lymphocytes, and is required for somatic hypermutation (SHM) and class-switch recombination of antibody genes (Honjo *et al.*, 2002). The hypothesis is based on the following observations: (1) AID can induce mutations in non-B cells (Yoshikawa *et al.*, 2002); (2) AID transgenic mice develop various tumours, including T-cell lymphoma and lung microadenoma (Okazaki *et al.*, 2003) and (3) AID can be induced in human hepatic, gastric and biliary epithelial cells when stimulated with pro-inflammatory cytokines, such as transforming growth factor-β and tumor necrosis factor-α, and when challenged with pathogens, such as hepatitis C virus (HCV) and *Helicobacter pylori*. AID is detected in the liver, stomach and bile duct in humans (Kou *et al.*, 2006; Endo *et al.*, 2007; Matsumoto *et al.*, 2007; Komori *et al.*, 2008). On the basis of this evidence, AID is a potential candidate for a mutagen in human cancers.

To substantiate this hypothesis, there is an urgent need for the establishment of an AID transgenic mouse model that recapitulates the development of human cancer. However, we have found earlier that it is difficult to analyse epithelial tumours of tissues other than lymphoid malignancies in a transgenic mouse model with constitutive and ubiquitous AID expression because of early death of the mice from lethal T-cell lymphoma (Okazaki *et al.*, 2003).

We generated earlier AID transgenic mice that can express AID conditionally in a Cre-recombianase-dependent manner (AID conditional transgenic, AID cTg) and reported B-cell-specific AID transgenic mice

Correspondence: Dr K Kinoshita, Evolutionary Medicine, Shiga Medical Center Research Institute, 5/4/1930, Moriyama, Shiga 524-8524, Japan.
E-mail: kkinoshi.shigamed@mac.com
Received 30 June 2008; revised 30 September 2008; accepted 14 October 2008; published online 10 November 2008

obtained by crossing AID cTg mice with B-cell-specific Cre transgenic mice (Muto *et al.*, 2006). In this study, we used a different Cre transgenic mouse line that expresses Cre in cells producing tissue nonspecific alkaline phosphatase (TNAP), as the mating partner of AID cTg mice.

TNAP is a member of the alkaline phosphatase family encoded by the *Akp2* gene and alternatively designated alkaline phosphatase 2, liver. As TNAP is also known as a marker of primordial germ cells and immature stem cells, including ES cells (Urven *et al.*, 1993; Kues *et al.*, 2005; Wang *et al.*, 2005), we initially speculated that AID expression in TNAP-positive cells might contribute to the development of tumours of germ cell origin. However, the resulting mice frequently developed hepatocellular carcinoma (HCC) but not germ cell tumours or lethal lymphomas. Histological and molecular analyses of HCC revealed genetic and phenotypic similarities to human tumours, including deleterious mutations in the p53 gene and the expression of α-fetoprotein, a well-known marker of human HCC. This mouse model should be useful for studies on the prevention and treatment of hepatic cancer, which is a major health concern worldwide.

Results

We crossed AID cTg mice with TNAP-Cre mice (Lomeli *et al.*, 2000; Figures 1a and b). AID cTg mice possess a CAG promoter-driven AID transgene, whose expression is blocked by insertion of the gene encoding enhanced green fluorescent protein (hereafter GFP) flanked by two loxP sites (Figure 1a). Therefore, the expression of Cre recombinase removes GFP and induces AID expression (Muto *et al.*, 2006). TNAP-Cre mice were generated by inserting the coding sequence of Cre recombinase into the *Akp2* locus of the 129/Sv mouse genome, and Cre was expressed in primordial germ cells in the embryonic genital ridge region (Lomeli *et al.*, 2000). We crossed homozygous female AID cTg mice on a C57BL/6 background with heterozygous male TNAP-Cre mice to obtain double-transgenic mice (TNAP-AID) and their littermate control AID cTg mice. Non-transgenic (wild-type, WT) and TNAP-Cre control mice on a C57BL/6:129/Sv hybrid background were obtained by mating female C57BL/6 mice with heterozygous male TNAP-Cre mice (Figure 1b). Accordingly, a total of 101 mice were prepared, which included 28 TNAP-AID, 29 AID cTg, 16 TNAP and 28 WT mice.

Most mice were viable at 90 weeks (Figure 1c), whereas in our earlier study, most AID transgenic mice died because of the lethal lymphoid malignancies, majority of which died within 50 weeks (Okazaki *et al.*, 2003). Although TNAP is expressed in primordial germ cells, we did not observe any incidence of testicular or ovarian tumours in TNAP-AID mice. Instead of germ cell tumours, we frequently observed macroscopic liver tumours in TNAP-AID mice. Table 1 summarizes

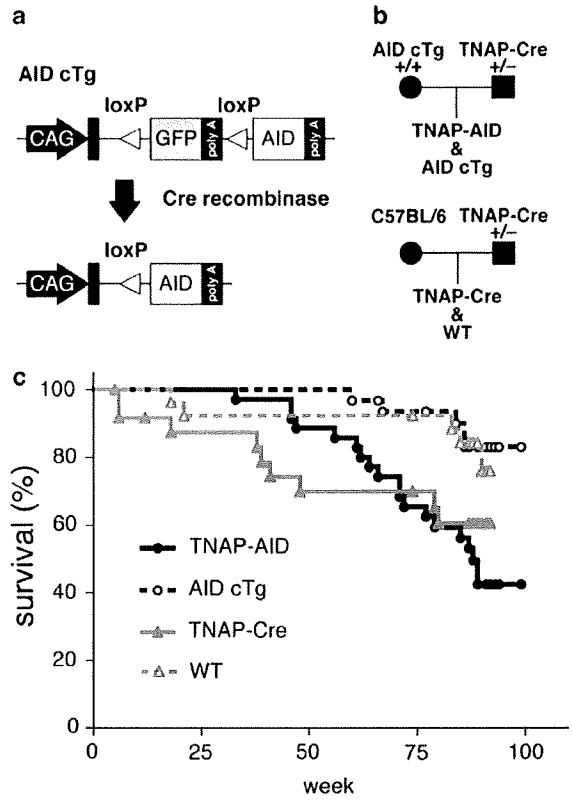

Figure 1 Generation of TNAP-AID mice and their controls. (**a**) Structure of transgene used to generate AID conditional transgenic mice (AID cTg) and the structure after Cre-mediated recombination. Arrows indicate the CAG promoter and rectangles indicate exons. Grey and open boxes represent the coding sequences of GFP and AID, respectively. A polyadenylation signal (poly A) is attached after each coding sequence. Triangles indicate loxP sites. (**b**) Mating scheme. For genotype abbreviations, see text. (**c**) Kaplan–Meier survival curves for the four genotype groups.

the frequency and spectrum of tumours that developed in TNAP-AID and control mice. Interestingly, liver tumours were observed in TNAP-AID mice but not in the control groups; the difference between TNAP-AID and AID cTg mice ($P=0.017$) and that between TNAP-AID and WT mice ($P=0.018$) was significant. There were no differences among the groups with regard to weights of body, liver, spleen and kidneys in healthy mice (data not shown).

Macroscopically, multiple liver nodules were frequently observed in mice with liver tumour (4 out of 15; Table 1, Figure 2a). The number and size of the nodules varied. Microscopically, liver tumours showed various degrees of cellular atypia, from adenoma to HCC with occasional 'carcinoma in adenoma', suggesting ongoing progression. The expression of α-fetoprotein, a marker of human HCC, was detected in the tumours by immunohistochemical analyses (Figure 2b) and quantitative reverse transcription PCR (qRT-PCR) (Supplementary Figure 1). None of the tumours showed

Table 1 Frequencies of tumours observed in TNAP-AID and control mice

Genotype	*Mean age at killing (weeks)*	*HCC*	*Lymphoma*	*Lung cancer*	*Stomach cancer*
TNAP-AID (15)	88.5	26.7% (4)	40.0% (6)	6.7% (1)	6.7% (1)
AID-cTg (24)	89.9	0.0% (0)	29.2% (7)	0.0% (0)	0.0% (0)
TNAP-Cre (14)	88.4	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
WT (23)	89.1	0.0% (0)	17.4% (4)	0.0% (0)	0.0% (0)

Abbreviations: HCC, hepatocellular carcinoma; WT, wild type.
Frequencies were calculated from the numbers of mice with macroscopic tumours of the indicated organs at approximately 90 weeks of age. Numbers in parentheses are the number of individuals killed for tumour inspection (genotype column) and those with macroscopic tumours (right four columns).

nuclear localization of β-catenin protein (data not shown).

Sequences of some tumour suppressor genes and oncogenes in the tumours were determined. Significant numbers of mutations were observed in the *Trp53* gene encoding the p53 protein (Table 2). The non-transcribed region upstream of the promoter did not contain mutations, an observation that is consistent with the transcription dependence of mutagenesis by AID (Yoshikawa *et al.*, 2002). Majority of mutations were single-base substitutions with significant GC bias, another footprint of AID (Figure 3a) (Yoshikawa *et al.*, 2002). There was no significant correlation between the distribution of mutations and those of SHM hot spot motifs (RGYW/WRCY) and DNA secondary structures (Supplementary Figure 2). Interestingly, the mutation frequency was the highest in cancer tissues followed by non-cancerous TNAP-AID liver and lowest in normal liver of WT mice, which corresponds to the PCR error rate or the background rate in this assay. The increase in mutation frequency in HCC compared with non-cancerous liver was because of the increase in single-base insertion events mostly at the homonucleotide tracts, leading to frameshift mutations in the 5′-half of the coding region (Table 2 and Figure 3b). Most of TNAP-AID liver mutations occurred in the DNA-binding domain of p53, which is critical for its function (Table 3).

Although small volumes (approximately 50 mm^3) of normal and cancerous liver tissues were dissected for mutation analysis, the identified mutations were quite diverse, suggesting that the sampled tissue block contained multiple HCC nodules, each representing monoclonal expansion. Such heterogeneity was verified by two independent series of PCR, plasmid cloning and sequencing for the same sample: there were few overlaps of mutations. For example, 8 and 7 clones derived from HCC of mouse TA113 contained nonsynonymous point mutations out of sequenced 27 and 40 clones in the first and second experiment, respectively. The mutation patterns were completely different except C808G (R270G), which appeared thrice and once in the first and the second trial, respectively. Table 3 lists the combined results of the two experiments. The most frequent mutations were C790T(R264W) and C808G(R270G), the latter corresponding to one of the second most important mutational targets (R273) in human cancers (Figure 3b, top). Despite extensive sequencing, no mutations were found in *Myc*, *Pten* and *Cdkn2a* (p16 and $p19^{Arf}$) genes (Supplementary Figure 3). The *Trp53* mutation frequency in the thymus was significantly lower than in the liver of TNAP-AID mice, which is consistent with the lack of lethal T-cell lymphomagenesis (Table 2).

To determine why liver tumours developed, TNAP expression in various organs was examined by qRT-PCR in E14.5 embryos and 53- to 74-week-old adult mice of strain C57BL/6. TNAP was ubiquitously expressed at a relatively high level in the liver and intestine of the embryos and in the liver and testis of the adults (Figure 4a). In agreement with this wide distribution of TNAP expression, GFP expression in TNAP-AID mice was absent in most organs in adults (Figure 4c). The substantial levels of AID expression were confirmed by qRT-PCR (Figure 4b) and western blotting (Figure 4d) in various organs of the adult TNAP-AID mice but not of AID cTg mice. The presence of AID protein and loss of GFP were detected in the non-cancerous liver sections of TNAP-AID mice (Figure 5). Taken together, these findings suggest that constitutive AID expression in hepatic lineage, including fetal and adult liver, contributes to the high incidence of liver tumours.

As reported for AID transgenic mice with the ubiquitous promoter, lymphomas were observed in TNAP-AID mice (Table 1). Although the frequency of lymphomas in the TNAP-AID group was higher than those in the other groups, the differences between TNAP-AID and AID cTg mice ($P=0.508$) and those between TNAP-AID and WT mice ($P=0.150$) were not statistically significant. The survival curve analysis (Figure 1c) suggested that there was no significant difference between the TNAP-AID and TNAP-Cre groups ($P=0.515$), indicating that the lymphomas in TNAP-AID mice were less aggressive than those in the transgenic mice with ubiquitous AID expression. Histologically, the lymphomas in TNAP-AID mice exhibited nodular appearance, which was different from that of the diffusely infiltrating lymphoma that develops in the ubiquitous promoter-driven AID transgenic mice (data not shown). This difference suggests that different mechanisms underlie tumorigenesis of lymphoid cells between these mice. We confirmed that GFP expression and Cre-mediated excision of GFP occurred efficiently in CD3-positive T cells of AID cTg and TNAP-AID mice, respectively (Supplementary Figure 4). This suggests that AID is expressed in T cells of TNAP-AID mice.

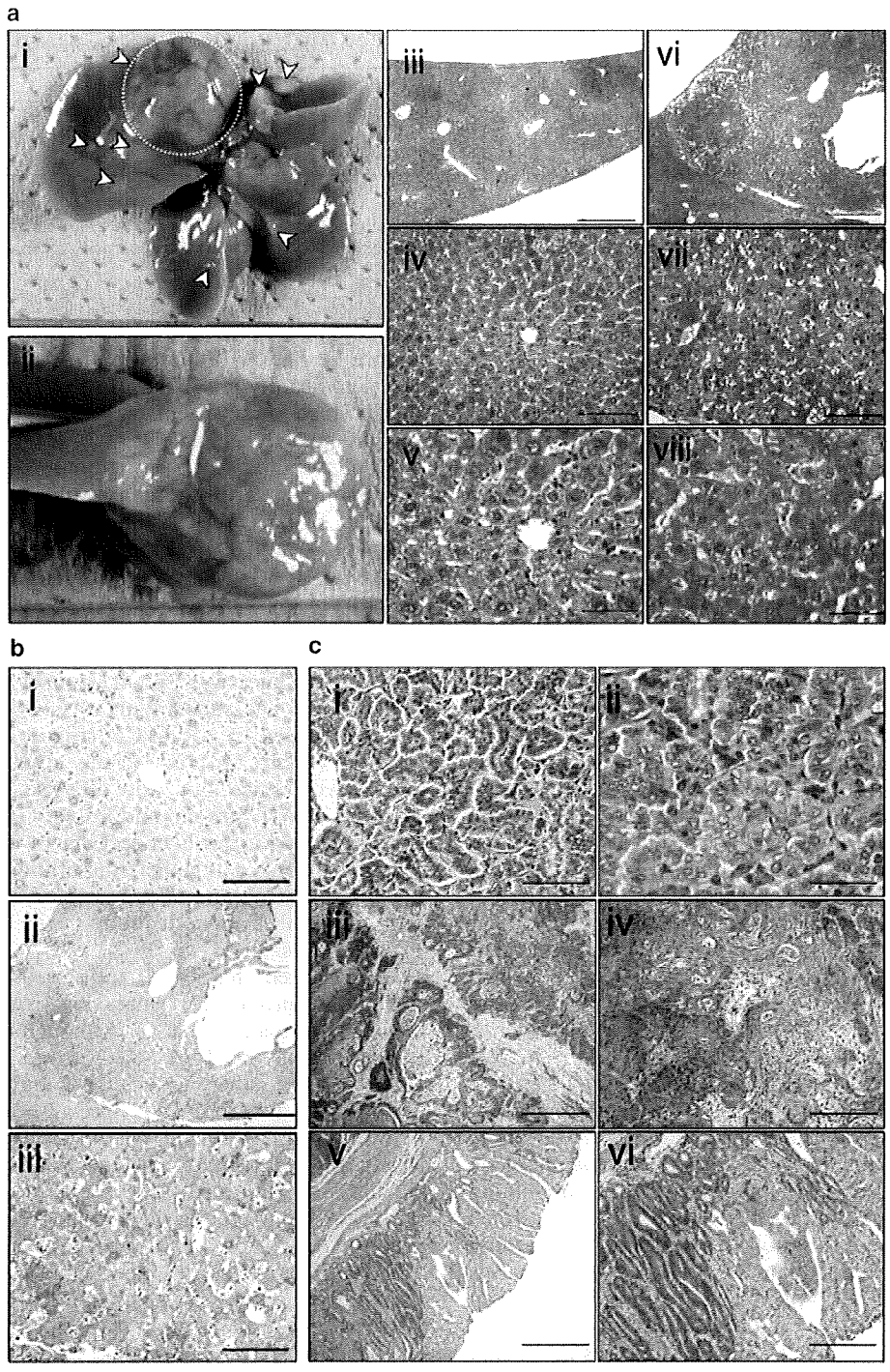

Figure 2 Tumours developed in TNAP-AID mice. (**a**) Macroscopic (i and ii) and microscopic (haematoxylin and eosin (HE) stain) images of HCC (i, ii, vi, vii and viii) that developed in a TNAP-AID mouse and non-cancerous liver of the same animal (iii, iv and v). In panel ai, arrowheads indicate liver tumour nodules. The largest nodule is encircled with a dotted line, and its cut surface is shown in panel aii). Scale bars are 1 mm (iii and iv), 500 μm (iv and vii) and 200 μm (v and viii). (**b**) Immunohistochemistry for α-fetoprotein of the normal liver (i) and HCC (ii and iii), which are serial sections of images shown in panel aiv, avi and aviii, respectively. Scale bars are 500 μm (i and iii) and 1 mm (ii). (**c**) Microscopic images (HE stain) of cancer of the lung (i and ii) and the stomach (iii and iv, squamous cell carcinoma; v and vi, adenocarcinoma). Scale bars are 200 μm (i, iv and vi), 100 μm (ii) and 500 μm (iii and v).

Table 2 *Trp53* gene mutation profiles in liver and HCC of TNAP-AID mice

Target	*Genotype*	*WT*	*TNAP-AID*	*HCC*	*Gastric cancer*	*Thymus*
	Tissue	*Normal liver*	*Non-cancerous liver*			
Transcribed region	Mutated/total clones	4/48	11/49	33/112	7/34	0/21
	Mutated (pm:ins:del)/total bases	4 (2:1:1)/52 487	12 (10:2:0)/55 321	42 (27:13:2)/127 674	9 (5:3:1)/38 789	0/24 446
	Mutation frequency $\times 10^4$	0.76	2.17	3.29	2.32	0.00
Non-transcribed region	Mutated/total clones	0/16	0/15	0/16		
	Mutated (pm:ins:del)/total bases	0/13 828	0/12 613	0/13 947		
	Mutation frequency $\times 10^4$	0.00	0.00	0.00		

Abbreviations: HCC, hepatocellular carcinoma; WT, wild type.
Mutation frequency of the transcribed and non-transcribed regions of the *Trp53* gene of the normal liver from two WT mice (87- and 89-week-old) and non-cancerous liver, HCC and thymic tissues from three TNAP-AID mice (92- to 94-week-old). Mutated base numbers are shown with numbers of point mutations (pm), insertions (ins) and deletions (del) in parentheses.

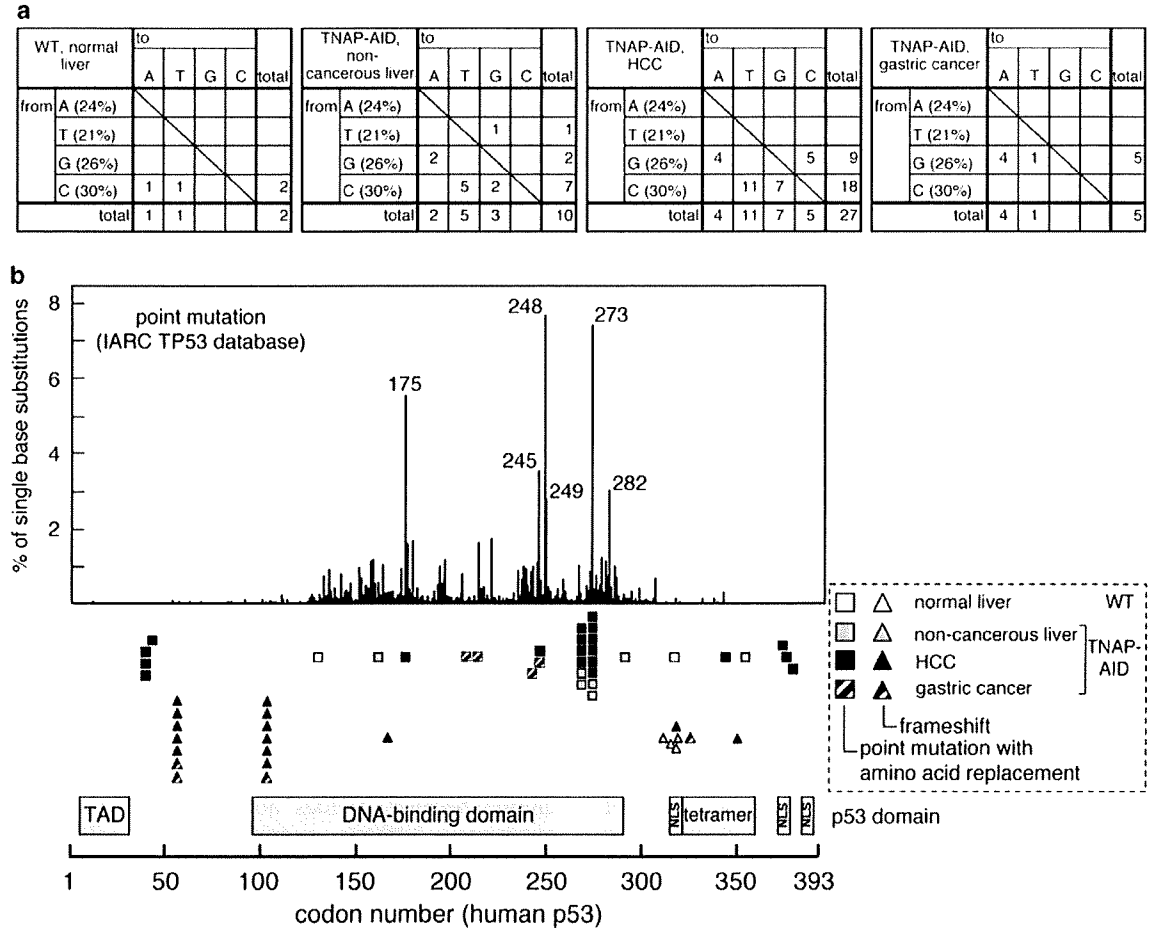

Figure 3 *Trp53* gene mutation profiles in liver and HCC of TNAP-AID mice. (**a**) Base substitution patterns seen in liver of WT mice (left), non-cancerous liver of TNAP-AID mice (centre) and HCC of TNAP-AID mice (right) extracted from the same data sets as those used for the mutation frequency analysis in Table 2. Percentages in parentheses are compositions of the indicated bases in the sequenced region. (**b**) Top: Distribution of *TP53* somatic mutations in human cancer, reproduced from the IARC TP53 Database (version R12), November 2007 (http://www-p53.iarc.fr/) (Petitjean *et al.*, 2007). Middle: Distribution of mouse *Trp53* mutations found in WT and TNAP-AID mouse liver. Codon positions are converted into human equivalents. Squares and triangles indicate point mutations with amino acid replacements and frameshifts, respectively. Open, grey, filled and hatched symbols indicate normal liver of WT mice, non-cancerous liver, HCC and gastric cancer of TNAP-AID mice, respectively. Bottom: p53 domain structure with transactivation (TAD), DNA-binding and tetramerization domains and nuclear localization signal (NLS).

Table 3 Predicted amino acid replacements observed in three TNAP-AID mice (TA113, TA114 and TA128), corresponding human mutations and affected domains of p53

Tissue	*Mutation*		*Mouse ID*	*Mutated/total clone*	*Human-equivalent codon*	*Domain*	*Functional assays*[a]		*Count*[a]
	Nucleotide	*Codon*					*Transactivation class*	*Structure-based prediction*	
TNAP-AID, non-cancerous liver	G472A	A158T	TA113	1/26	A161T	DNA binding	Partially functional	Non-functional	73
	C790T	R264W	TA128	2/23	R267W	DNA binding	Non-functional	Non-functional	25
	C808G	R270G	TA128	2/23	R273G	DNA binding	Non-functional	Non-functional	14
	G860A	R287H	TA113	1/26	R290H	DNA binding	Supertrans	NA	29
	C1049T	A350V	TA113	1/26	A353V	Tetramerization	Supertrans	NA	0
TNAP-AID, HCC	G119C	C40S	TA113	2/67	None	Transactivation	NA	NA	NA
			TA114	1/33					
	G127A	D43N	TA113	1/67	D42N	Transactivation	Functional	NA	0
	C532T	R178C	TA114	1/33	R175C	DNA binding	Partially functional	Non-functional	24
	G729A	M243I	TA113	1/67	M246I	DNA binding	Non-functional	Functional	32
	C790T	R264W	TA113	3/67	R267W	DNA binding	Non-functional	Non-functional	25
			TA114	1/33					
	C808G	R270G	TA113	4/67	R273G	DNA binding	Non-functional	Non-functional	14
			TA114	2/33					
	G1020C	E340D	TA113	1/67	E343D	Tetramerization	Functional	NA	0
	G1112A	G371D	TA113	1/67	G374D	Regulation/NLS	Functional	NA	0
	C1118G	S373C	TA113	1/67	S376C	Regulation/NLS	Functional	NA	0
	G1127C	R376P	TA113	1/67	R379P	Regulation	Partially functional	NA	0
TNAP-AID, gastric cancer	G613T	D205Y	TA114	1/34	D208Y	DNA binding	Non-functional	Functional	1
	G629A	R210H	TA114	1/34	R213Q	DNA binding	Non-functional	Non-functional	34
	G716A	C239Y	TA114	1/34	C242Y	DNA binding	Non-functional	Non-functional	51
	G729A	M243I	TA114	1/34	M246I	DNA binding	Non-functional	Functional	32

Abbreviation: NA, data not available.
For details, visit the IARC website.
[a]Based on the data presented in the IARC website and described in the legend of Figure 3. Functional assays are predicted as impacts of mutations in yeast assays and structural computation. Count is the number of the entries in the database.

Macroscopic tumours of other organs included one case of lung adenocarcinoma and one case of gastric cancer (squamous cell carcinoma) (Table 1 and Figure 2c). These rare tumours were restricted to the TNAP-AID group and likely to be caused by AID expression. However, the small number of tumours precludes a quantitative discussion on causal relationships, except for liver tumours. Sequencing analysis revealed that *Trp53* mutation frequency in gastric cancer tissue was $2.32/10^4$, a comparable level to that observed in HCC tissues (Tables 2 and 3 as well as Figure 3). Unfortunately, we could not carry out mutation analyses on the lung cancer tissue because of its small size. We did not examine microscopic tumours except for subpleural lung microadenomas, which were observed in the TNAP-AID mice, a finding similar to that in the AID transgenic mice in our earlier study (Okazaki *et al.*, 2003). In addition, two cases of gastric adenocarcinoma were found incidentally by microscopic examination (Figure 2c), suggesting that many microscopic tumours may have been overlooked.

Discussion

Here we describe TNAP-AID mice as a novel model of hepatocarcinogenesis, which occurs as a result of the accumulation of mutations in chronically inflamed liver in humans, often caused by infection with hepatitis viruses. The TNAP-AID mouse model exhibits several characteristics of human HCC, in that the mice develop HCC spontaneously and the HCC tissue expresses α-fetoprotein and that it has the p53 gene mutations, some of which cause the same amino acid replacements as those seen in human tumours. Earlier HCC mouse models include mice with genetic modifications of *Lkb1*, *Mdr2*, *Aox* and *Pten* genes (Fan *et al.*, 1998; Nakau *et al.*, 2002; Horie *et al.*, 2004; Katzenellenbogen *et al.*, 2006), transgenic mice overexpressing c-myc, transforming growth factor-α, transforming growth factor-β1, HBx of hepatitis B virus and HCV core (Sandgren *et al.*, 1989; Kim *et al.*, 1991; Murakami *et al.*, 1993; Koike *et al.*, 1994, 2002; Factor *et al.*, 1997; Riehle *et al.*, 2008) and chemical- or diet-induced HCC (Sell, 2001; Maeda *et al.*, 2005; Ma *et al.*, 2006; Sakurai *et al.*, 2006). In contrast to these models, our TNAP-AID model is unique because it does not arbitrarily target specific oncogenes, tumour suppressors or stability genes. In addition, it does not use chemical mutagens or specific dietary conditions that are not associated with human HCC. The development of *Trp53* mutations is another feature of this model because it is unprecedented in any earlier mouse HCC models as mentioned above.

Liver tumours were occasionally observed in our earlier AID transgenic mice (Endo *et al.*, 2007). However, it was difficult to use these mice as a liver tumour model because of early death from lethal lymphomas. Trials to detect *Trp53* mutations in the liver tumour failed (data not shown). Even the T-cell

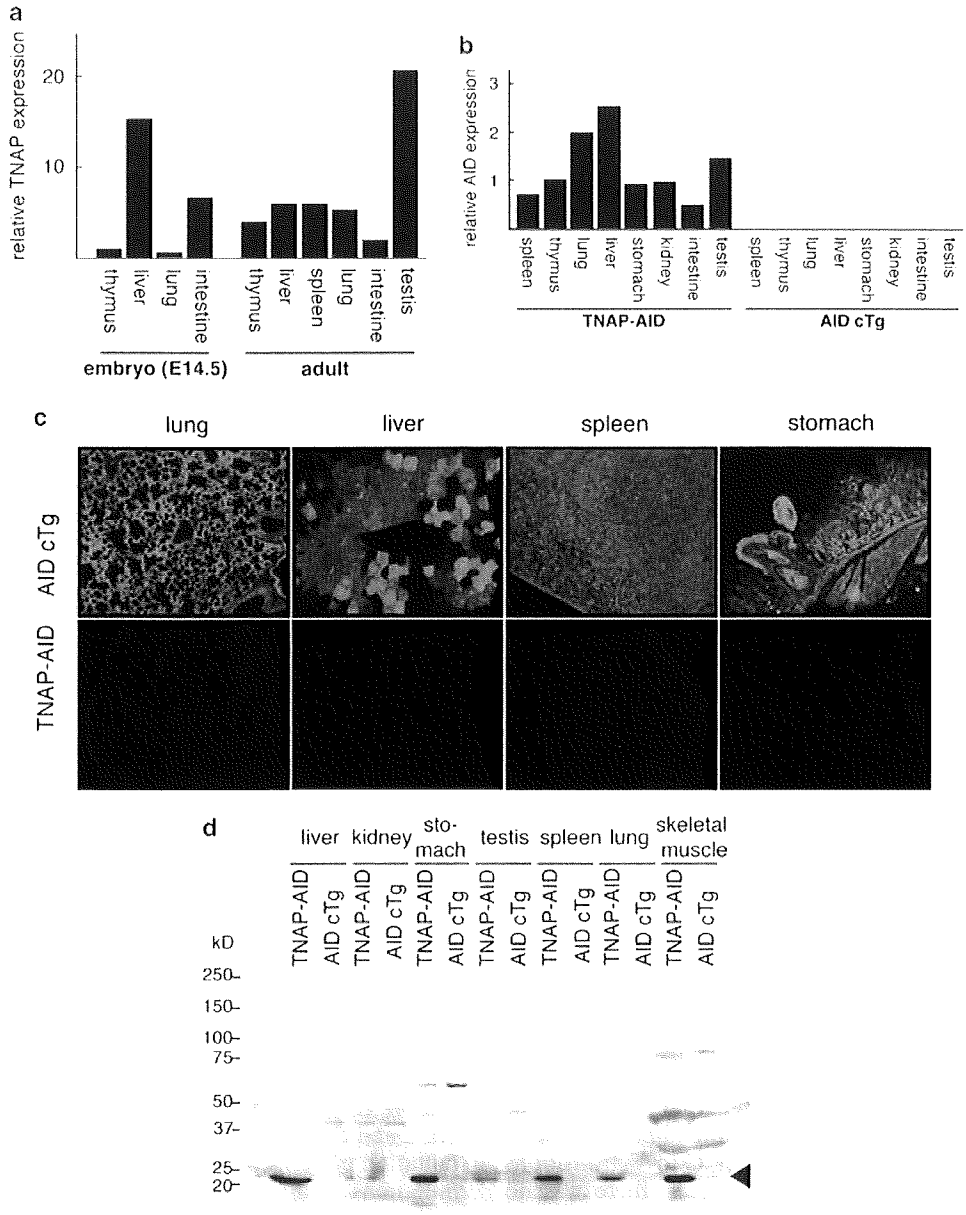

Figure 4 Expression analysis of TNAP-AID mice. (**a**) Results of qRT-PCR for TNAP gene (*Akp2*) calibrated by the amount of ribosomal 18S RNA for indicated organs of adult mice (53- and 74-week-old) and four E14.5 embryos. The graph shows average results for the indicated groups. (**b**) Results of qRT-PCR for transgene-derived AID expression. (**c**) GFP fluorescence in frozen sections of GFP from the indicated organs and animals. A mosaic pattern of GFP expression in the liver of AID cTg mice was always observed, which may be attributed to the integration site of the transgene. (**d**) Western blot analysis using anti-AID monoclonal antibody MAID-1 for the lysates of indicated organs of 42-week-old TNAP-AID and AID cTg mice. Arrowhead indicates the position of the AID band. Comparable intensities of nonspecific bands indicate equal loading of extracts between TNAP-AID and AID cTg lanes.

lymphomas that developed in the earlier AID transgenic mice did not harbour *Trp53* mutations (Kotani *et al.*, 2005). Therefore, the TNAP-AID mouse model is the first model in which *Trp53* mutations can be detected unequivocally in non-cancerous tissues as well as in cancerous tissues that express AID. In spite of these differences, the two independent AID transgenic mice lines commonly developed liver tumours, making it unlikely that liver tumours in TNAP-AID mice are ascribed to unexpected variables such as the transgene integration sites and the segregation of relevant loci upon crossing of mice with different genetic backgrounds. However, the possibility that these variables somewhat modified the tumour frequency cannot be excluded.

The overlap of mutational hot spots between TNAP-AID and human HCC (Table 3 and Figure 3b) suggests that these mutations induced by AID have critical roles.

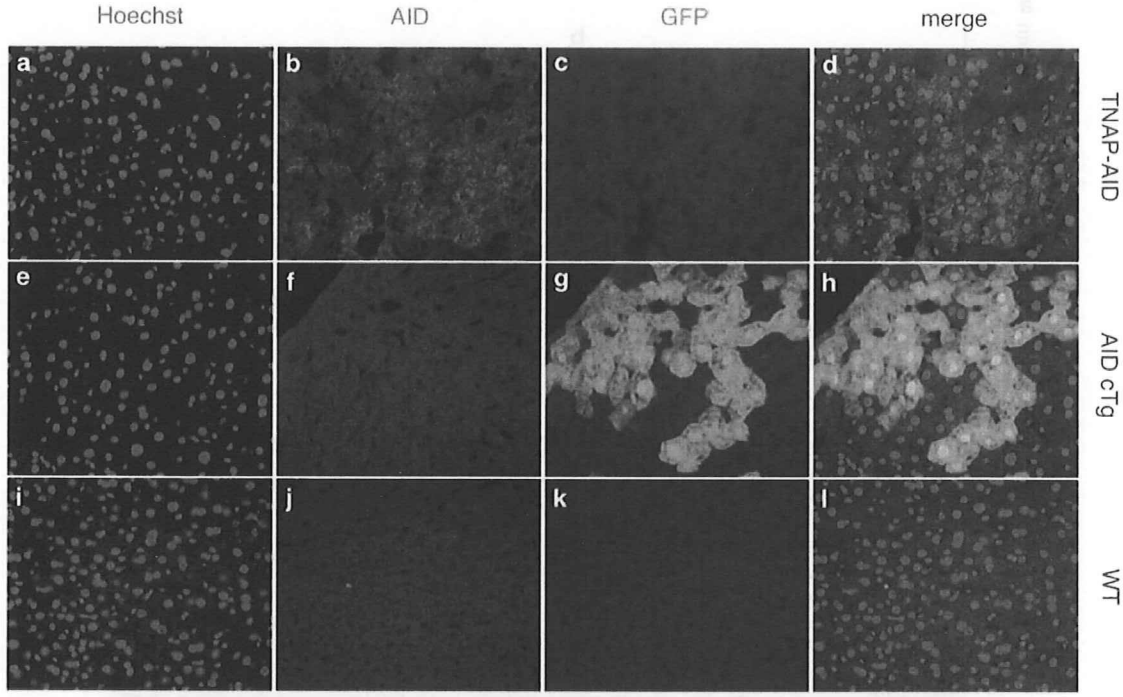

Figure 5 Immunohistochemical analysis of frozen liver sections from TNAP-AID (**a–d**), AID cTg (**e–h**) and WT (**i–l**) mice. (**a**, **e** and **i**) Hoechst 33342 staining; (**b**, **f** and **j**) AID immunostaining; (**c**, **g** and **k**) GFP fluorescence; (**d**, **h** and **l**) merge of the three colours.

The lack of *Trp53* mutations in the earlier transgenic mouse model may reflect a different evolutionary path from the initiation to progression of liver tumours between the two models despite the common involvement of AID. This difference may be related to the difference in the frequency and aggressiveness of lymphomas between TNAP-AID and the earlier AID transgenic mouse, as discussed below. The point mutations observed in *Trp53* are likely to be a direct effect of AID activity, because they share some molecular footprints with SHM of the immunoglobulin genes, including GC bias, frequent C-to-T transitions and transcription dependence. However, there was no correlation between the distributions of mutations and SHM hot spot motifs or DNA secondary structures in the *Trp53* sequence. As these parameters show only weak association even for SHM, the small numbers of point mutations observed in this study may not be sufficient to achieve statistical significance. On the other hand, single-base insertions causing frameshifts in *Trp53* were frequently observed in HCC but not in the non-cancerous liver of TNAP-AID mice. As single-base insertions are relatively rare in SHM induced by AID, these might be caused by genetic instability (for example, impaired mismatch repair) acquired during the progression of the tumours.

The presence of mutations in non-cancerous tissue expressing AID is reminiscent of a study on human patients with chronic hepatitis (Kou *et al.*, 2006), together suggesting a causal role of AID in tumorigenesis. It is noteworthy that the mutations corresponding to two mutational hot spots in *TP53* gene encoding human p53 (R273 and R175) were included in mutations found in TNAP-AID HCC. Other mutations were mapped close to the human hot spots (M246 and R267) (Figure 3b and Table 3). Among them three patterns of amino acid replacement, C40S, R264W and R270G, were observed repeatedly in multiple individuals. Such repeated appearance of the specific mutations in different individuals may indicate their significance in tumorigenesis. In addition to these common mutations, HCC and the non-cancerous liver contained quite diverse mutation patterns, which parallels heterogeneity of *TP53* mutations within human HCC occurred in a single individual and is consistent with the 'field cancerization' model (Thorgeirsson and Grisham, 2002).

The utility of the TNAP-AID mouse as a human HCC model depends partly on the lack of lethal lymphomas that were observed in almost all AID transgenic mice in our earlier study. To examine the reason why lymphomas were infrequent and less aggressive in the TNAP-AID model, mutations caused by AID and its expression in T lymphocytes of TNAP-AID mice were analysed by sequencing the *Trp53* coding sequence (Table 2) and flow cytometry (Supplementary Figure 4). From these analyses, it was found that *Trp53* was not mutated, but AID was expressed in the thymocytes of the TNAP-AID mice. The level of AID expression in the thymus of TNAP-AID mice was half than that seen in the earlier AID transgenic line in a qRT-PCR analysis (data not shown). We speculate that the difference in the levels of AID expression in T cells

may be one of the reasons for this, although the difference in the genetic background (C57BL/6:129/Sv hybrid versus pure C57BL/6) cannot be excluded.

The TNAP-AID mice did not develop germ cell tumours despite the expression of AID in the testes (Figure 4d) and the absence of GFP in the seminiferous tubules and oocytes of the ovaries (data not shown). We reported earlier the lack of lymphomagenesis in B-cell-specific AID transgenic mice obtained by crossing the same AID cTg line with CD19-Cre mice (Muto *et al.*, 2006), suggesting that some protective mechanism might exist in B cells, which have physiological expression of AID. A similar reason might apply to germ cells because AID has been reported to be expressed physiologically in the human testis (Schreck *et al.*, 2006) and mouse ovary (Morgan *et al.*, 2004).

To conclude, we generated a new line of AID transgenic mice that develop HCC but not lethal T-cell lymphoma. We consider that the TNAP-AID mouse model is useful for human HCC studies because it has been shown that AID is induced in the human pre-cancerous conditions of chronic hepatitis and cirrhosis (Kou *et al.*, 2006). The relatively long latency before HCC development in this model is reasonable, considering that it is a physiological recapitulation of a human tumour phenotype that takes decades to develop.

Materials and methods

Mice

The AID cTg mice (line 20) on a C57BL/6 background (Muto *et al.*, 2006) were self-crossed to obtain homozygous transgenic mice, which were maintained in a specific pathogen-free facility at the Kyoto University Faculty of Medicine and Shiga Medical Center. This mouse line has been deposited at the Riken Bioresource Center (Tsukuba, Japan; No. RBRC00892). The TNAP-Cre mice (Lomeli *et al.*, 2000) were gifted by Dr Andras Nagy (Mount Sinai Hospital, Toronto, Canada) and maintained by self-crossing between the heterozygous mice. C57BL/6 mice were purchased from Shimizu Laboratory Supplies Co., Ltd (Kyoto, Japan). All mice were fed *ad libitum* and were killed by cervical dislocation for censoring, or observed until spontaneous death. Upon censoring and spontaneous death, the numbers of macroscopic tumours were counted after laparotomy and thoracotomy. Kaplan–Meier survival curves were analysed using Prism 4.0 software (Graphpad, San Diego, CA, USA). All animal experiments were approved by the Ethical Committee for Animal Experiments and performed under the Guidelines for Animal Experiments of Kyoto University and Shiga Medical Center.

Histology and immunohistochemistry

Paraffin-embedded mouse organs were sectioned and stained with haematoxylin and eosin. α-Fetoprotein was detected using anti-α-fetoprotein antibody (Dako, Glostrup, Denmark) and an ABC kit (Vector, Burlingame, CA, USA) for paraffin-embedded sections of formalin-fixed liver. For AID immunostaining, freshly isolated liver was fixed in 4% paraformaldehyde in phosphate-buffered saline on ice for 2 h and immersed in 30% sucrose for 18 h. After embedding in OCT compound (Sakura Fineteck Japan, Tokyo, Japan) in liquid nitrogen, 8-μm sections were sliced and mounted on glass slides. Hoechst 33342 dye was used to visualize nuclei. AID protein was detected using the monoclonal anti-AID antibody, MAID-2 (Tsuji *et al.*, 2008), with peroxidase-labelled donkey F(ab')2 anti-rat IgG (Jackson ImmunoResearch, West Grove, PA, USA) and signal amplification using TSA Plus DNP and TSA Plus TMR kits (Perkin Elmer, Waltham, MA, USA). Images taken with a fluorescence microscope (DM5000B; Leica, Wetzlar, Germany) were merged using Photoshop (Adobe, San Jose, CA, USA).

Mutation analysis

Cancerous and non-cancerous liver tissues of approximate volume of 50 mm^3 were macroscopically dissected, frozen in liquid nitrogen and powdered with a frozen-cell crusher (Cryo-press; Microtec, Funabashi, Japan). Genomic DNA and total RNA were extracted using the QIAamp DNA Mini kit (Qiagen, Duesseldorf, Germany) and the QuickGene kit (Fuji, Tokyo, Japan) from liquid nitrogen-frozen tissues that had been dissected macroscopically from non-cancerous livers and cancerous nodules. Sequencing of *Trp53* was performed as described earlier (Matsumoto *et al.*, 2007) except for primers for the non-transcribed regions, sequences of which are shown in Supplementary Figure 5. Mutations of the *Myc*, *Pten* and *Cdkn2a* genes were analysed as described (Matsumoto *et al.*, 2007) except for primers, sequences of which are shown in Supplementary Figure 5. Statistical significance was assessed by Fisher's exact test using Stata 8.2 software (Stata Corp, College Station, TX, USA). Searching for SHM hot spot motifs and DNA secondary structure analyses were performed using GENETYX-MAC 14 (Genetyx Corp., Tokyo, Japan) and mfold software (Zuker, 2003), respectively.

qRT-PCR

cDNA was synthesized using the iScript kit (Bio-Rad, Hercules, CA, USA). Quantitative PCR (qPCR) was performed using QuantiTect reagent (Qiagen) and a real-time thermal cycler (Mx3000P; Stratagene, La Jolla, CA, USA). Oligonucleotide sequences are shown in Supplementary Figure 5.

Western blotting

Mouse organs were dissected, frozen in liquid nitrogen and powdered with a frozen-cell crusher (Cryo-press). Proteins were extracted from the tissue powder using 10 mM sodium phosphate (pH 6.8), 200 mM NaCl, 1.5 mM $MgCl_2$ and 0.2 mM EDTA with a protease inhibitor cocktail (Complete; Roche Diagnostics, Mannheim, Germany). Protein of 100 μg per lane was applied to a 5–20% polyacrylamide gel (Bio-Rad), electrophoresed and blotted on to Hybond P membrane (GE Healthcare, Buckinghamshire, UK). Signals were generated using peroxidase-labelled anti-rat IgG (Jackson ImmunoResearch) and the ECL Plus system (GE Healthcare) and detected with a LAS-3000 mini image analysis system (Fuji).

Acknowledgements

We thank Dr Andras Nagy for his generous gift of TNAP-Cre mice, Dr Takashi Shinohara for suggesting the choice of the Cre mouse strain, Dr Yoshinobu Toda for the preparation of tissue sections and Dr Masayuki Tsuji for technical help with the immunohistochemical analyses. We also thank Dr Sidonia Fagarasan for critical reading of the manuscript and discussions. This study was supported by Grant-in-Aid for Scientific Research (17013042 and 18390122) and the Takeda Science Foundation.

References

Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM *et al.* (2007). Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene* **26**: 5587–5595.

Factor VM, Kao CY, Santoni-Rugiu E, Woitach JT, Jensen MR, Thorgeirsson SS. (1997). Constitutive expression of mature transforming growth factor beta1 in the liver accelerates hepatocarcinogenesis in transgenic mice. *Cancer Res* **57**: 2089–2095.

Fan CY, Pan J, Usuda N, Yeldandi AV, Rao MS, Reddy JK. (1998). Steatohepatitis, spontaneous peroxisome proliferation and liver tumors in mice lacking peroxisomal fatty acyl-CoA oxidase. Implications for peroxisome proliferator-activated receptor alpha natural ligand metabolism. *J Biol Chem* **273**: 15639–15645.

Honjo T, Kinoshita K, Muramatsu M. (2002). Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu Rev Immunol* **20**: 165–196.

Horie Y, Suzuki A, Kataoka E, Sasaki T, Hamada K, Sasaki J *et al.* (2004). Hepatocyte-specific Pten deficiency results in steatohepatitis and hepatocellular carcinomas. *J Clin Invest* **113**: 1774–1783.

Katzenellenbogen M, Pappo O, Barash H, Klopstock N, Mizrahi L, Olam D *et al.* (2006). Multiple adaptive mechanisms to chronic liver disease revealed at early stages of liver carcinogenesis in the Mdr2-knockout mice. *Cancer Res* **66**: 4001–4010.

Kim CM, Koike K, Saito I, Miyamura T, Jay G. (1991). HBx gene of hepatitis B virus induces liver cancer in transgenic mice. *Nature* **351**: 317–320.

Kinoshita K, Nonaka T. (2006). The dark side of activation-induced cytidine deaminase: relationship with leukemia and beyond. *Int J Hematol* **83**: 201–207.

Koike K, Moriya K, Iino S, Yotsuyanagi H, Endo Y, Miyamura T *et al.* (1994). High-level expression of hepatitis B virus HBx gene and hepatocarcinogenesis in transgenic mice. *Hepatology* **19**: 810–819.

Koike K, Moriya K, Kimura S. (2002). Role of hepatitis C virus in the development of hepatocellular carcinoma: transgenic approach to viral hepatocarcinogenesis. *J Gastroenterol Hepatol* **17**: 394–400.

Komori J, Marusawa H, Machimoto T, Endo Y, Kinoshita K, Kou T *et al.* (2008). Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology* **47**: 888–896.

Kotani A, Okazaki IM, Muramatsu M, Kinoshita K, Begum NA, Nakajima T *et al.* (2005). A target selection of somatic hypermutations is regulated similarly between T and B cells upon activation-induced cytidine deaminase expression. *Proc Natl Acad Sci USA* **102**: 4506–4511.

Kou T, Marusawa H, Kinoshita K, Endo Y, Okazaki IM, Ueda Y *et al.* (2006). Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer* **120**: 469–476.

Kues WA, Petersen B, Mysegades W, Carnwath JW, Niemann H. (2005). Isolation of murine and porcine fetal stem cells from somatic tissue. *Biol Reprod* **72**: 1020–1028.

Lomeli H, Ramos-Mejia V, Gertsenstein M, Lobe CG, Nagy A. (2000). Targeted insertion of Cre recombinase into the TNAP gene: excision in primordial germ cells. *Genesis* **26**: 116–117.

Ma W, Xia X, Stafford LJ, Yu C, Wang F, LeSage G *et al.* (2006). Expression of GCIP in transgenic mice decreases susceptibility to chemical hepatocarcinogenesis. *Oncogene* **25**: 4207–4216.

Maeda S, Kamata H, Luo JL, Leffert H, Karin M. (2005). IKKbeta couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* **121**: 977–990.

Marusawa H. (2008). Aberrant AID expression and human cancer development. *Int J Biochem Cell Biol* **40**: 1399–1402.

Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T *et al.* (2007). Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med* **13**: 470–476.

Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**: 52353–52360.

Murakami H, Sanderson ND, Nagy P, Marino PA, Merlino G, Thorgeirsson SS. (1993). Transgenic mouse model for synergistic effects of nuclear oncogenes and growth factors in tumorigenesis: interaction of c-myc and transforming growth factor alpha in hepatic oncogenesis. *Cancer Res* **53**: 1719–1723.

Muto T, Okazaki IM, Yamada S, Tanaka Y, Kinoshita K, Muramatsu M *et al.* (2006). Negative regulation of activation-induced cytidine deaminase in B cells. *Proc Natl Acad Sci USA* **103**: 2752–2757.

Nakau M, Miyoshi H, Seldin MF, Imamura M, Oshima M, Taketo MM. (2002). Hepatocellular carcinoma caused by loss of heterozygosity in Lkb1 gene knockout mice. *Cancer Res* **62**: 4549–4553.

Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K *et al.* (2003). Constitutive expression of AID leads to tumorigenesis. *J Exp Med* **197**: 1173–1181.

Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P *et al.* (2007). Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* **28**: 622–629.

Riehle KJ, Campbell JS, McMahan RS, Johnson MM, Beyer RP, Bammler TK *et al.* (2008). Regulation of liver regeneration and hepatocarcinogenesis by suppressor of cytokine signaling 3. *J Exp Med* **205**: 91–103.

Sakurai T, Maeda S, Chang L, Karin M. (2006). Loss of hepatic NF-kappa B activity enhances chemical hepatocarcinogenesis through sustained c-Jun N-terminal kinase 1 activation. *Proc Natl Acad Sci USA* **103**: 10544–10551.

Sandgren EP, Quaife CJ, Pinkert CA, Palmiter RD, Brinster RL. (1989). Oncogene-induced liver neoplasia in transgenic mice. *Oncogene* **4**: 715–724.

Schreck S, Buettner M, Kremmer E, Bogdan M, Herbst H, Niedobitek G. (2006). Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. *J Pathol* **210**: 26–31.

Sell S. (2001). Heterogeneity and plasticity of hepatocyte lineage cells. *Hepatology* **33**: 738–750.

Thorgeirsson SS, Grisham JW. (2002). Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet* **31**: 339–346.

Tsuji M, Suzuki K, Kitamura H, Maruya M, Kinoshita K, Ivanov II *et al.* (2008). Requirement for lymphoid tissue-inducer cells in isolated follicle formation and T cell-independent immunoglobulin a generation in the gut. *Immunity* **29**: 261–271.

Urven LE, Weng DE, Schumaker AL, Gearhart JD, McCarrey JR. (1993). Differential gene expression in fetal mouse germ cells. *Biol Reprod* **48**: 564–574.

Vogelstein B, Kinzler KW. (2004). Cancer genes and the pathways they control. *Nat Med* **10**: 789–799.

Wang L, Duan E, Sung LY, Jeong BS, Yang X, Tian XC. (2005). Generation and characterization of pluripotent stem cells from cloned bovine embryos. *Biol Reprod* **73**: 149–155.

Yoshikawa K, Okazaki IM, Eto T, Kinoshita K, Muramatsu M, Nagaoka H *et al.* (2002). AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296**: 2033–2036.

Zuker M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)