Materials and methods

Patients

From 2007 to 2012, samples for the GWAS were obtained from 416 CHC patients who were treated at 22 hospitals (liver units with hepatologists) throughout Japan. In the following stage of replication analysis, samples were collected in an independent set of 404 Japanese CHC patients. Most patients were treated with PEG-IFN-α2b (1.5 μg/kg body weight subcutaneously once a week) or PEG-IFN-α2a (180 μg once a week) plus RBV (600–1,000 mg daily according to body weight) for 48 weeks for HCV genotype 1 and 24 weeks for genotype 2. Treatment duration was extended in some patients up to 72 weeks for genotype 1 and 48 weeks for genotype 2 according to physicians' preferences. Other patients were treated with PEG-IFN-α2a or IFN monotherapy, or IFN-α2b plus RBV in standard doses of the regimens. The doses of drugs were reduced according to the recommendations on the package inserts or the clinical conditions of the individual patients. Erythropoietin or other growth factors were not given. Patients chronically infected with hepatitis B virus or human immunodeficiency virus, or with other causes of liver disease such as autoimmune hepatitis and primary biliary cirrhosis, were excluded from this study. Written informed consent was obtained from all individual participants in this study and the study protocol conformed to the ethics guidelines of the Declaration of Helsinki and was approved by the institutional ethics review committees.

Inclusion criteria of neutropenia

In the initial stage of GWAS, we defined the inclusion criteria of the case group as minimum neutrophil counts of $<750/mm^3$ at week 2 or 4 during IFN-based therapy, since the dose reduction of IFN is recommended at those levels on the package inserts. Thereafter we did it as minimum neutrophil counts of $<600/mm^3$ at week 2 or 4 in the following GWAS and the replication stages.

SNP genotyping and data cleaning

We conducted two stages of GWAS using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Inc. Santa Clara, CA) according to the manufacturer's instructions. The cut-off value was calculated to maximize the difference, which was also close to median change. At GWAS, the average overall call rate of patients in the case and the control group reached 98.66 and 98.79 %, respectively. We then applied the following thresholds for SNP quality control (QC) in data cleaning: SNP call rate ≥95 % for all samples, minor allele frequency (MAF) ≥1 % for all samples. A total of 601,578 SNPs on autosomal chromosomes passed the QC filters and were used for association analysis. All cluster plots of SNPs showing $P < 0.0001$ in association analyses by comparing allele frequencies in both groups were checked by visual inspection and SNPs with ambiguous genotype calls were excluded. In the replication study, the genotyping of 192 candidate SNPs in an independent set of 404 Japanese HCV-infected patients was carried out using the DigiTag2 assay (Nishida et al. 2007). Successfully genotyped SNPs in the replication analysis had a >95 % call rate, and cleared Hardy–Weinberg equilibrium (HWE) $P \geq 0.001$. One SNP could not be genotyped, and hence we obtained data on 191 SNPs including rs9915252. Three SNPs, rs4794822, rs3907022, and rs3859192 located around the proteasome 26S subunits non-ATPase 3 (*PSMD3*) gene and rs8099917 near the *IL28B* gene were genotyped by TaqMan SNP Genotyping Assays (Applied Biosystems, Carlsbad, CA) following the manufacturer's protocol.

Laboratory and histological tests

Blood samples were obtained at baseline and at appropriate periods after the start of therapy and for hematologic tests, blood chemistry, and HCV RNA. Fibrosis was evaluated on a scale of 0–4 according to the METAVIR scoring system. The SVR was defined as an undetectable HCV RNA level by Roche COBAS Amplicor HCV Monitor test, v.2.0 (Roche Molecular Diagnostics, Pleasanton, CA) with a lower detection limit of 50 IU/ml or Roche COBAS AmpliPrep/COBAS TaqMan HCV assay (Roche Molecular Diagnostics, Pleasanton, CA) with a lower detection limit of 15 IU/ml 24 weeks after the completion of therapy. Serum granulocyte colony-stimulating factor (G-CSF) levels were analyzed using Human G-CSF Quantikine ELISA Kit (R&D Systems, Inc., Minneapolis, MN).

K. Yamamoto
Department of Gastroenterology and Hepatology, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences, Okayama 700-0914, Japan

H. Watanabe
Department of Gastroenterology, Yamagata University Faculty of Medicine, Yamagata 990-9585, Japan

S. Hige
Department of Gastroenterology, Sapporo Kosei General Hospital, Sapporo 060-0033, Japan

A. Matsumoto · E. Tanaka
Department of Medicine, Shinshu University School of Medicine, Matsumoto 390-8621, Japan

Expression quantitative trait locus analysis

Expression quantitative trait locus analysis (eQTL) was conducted using the web-based tool, Genevar (http://www.sanger.ac.uk/resources/software/genevar) (Yang et al. 2010). We evaluated the correlations between rs2305482 genotypes and the expression of transcripts of *PSMD3* or colony-stimulating factor 3 (*CSF3*) by the Spearman's rank correlation coefficient.

Statistical analysis

In the GWAS and the replication stages, the observed association between a SNP and neutropenia induced by IFN-based therapy was assessed by the Chi square test with a two-by-two contingency table in three genetic models: the allele frequency model, the dominant-effect model and the recessive-effect model. Significance levels after Bonferroni correction for multiple testing were $P = 8.31 \times 10^{-8}$ (0.05/601,578) in the GWAS stage and $P = 2.62 \times 10^{-4}$ (0.05/191) in the replication stage. Categorical variables were compared between groups by the Chi square test, and non-categorical variables by the Student's t test or the Mann–Whitney U test. Multivariate logistic regression analysis with stepwise forward selection was performed with $P < 0.05$ in univariate analysis as the criteria for model inclusion. To evaluate the discriminatory ability of neutrophil counts at baseline to predict neutropenia during IFN-based therapy, receiver operating characteristic curve (ROC) curve analysis was conducted. Changes of serum G-CSF levels from baseline to the period with neutropenia during IFN-based therapy were compared by the repeated measure analysis of variance (ANOVA). Correlations between neutrophil counts and serum G-CSF levels were analyzed using Pearson's correlation coefficient test. $P < 0.05$ was considered significant in all tests.

Results

Genetic variants associated with IFN-induced neutropenia

We conducted two stages of GWAS by changing the terms of neutrophil counts, followed by the replication analysis (Fig. 1). The characteristics of the patients in each group for the GWAS and the replication stage are summarized in Table 1. At the first stage of GWAS (GWAS-1st), we genotyped 416 Japanese CHC patients with minimum neutrophil counts of <750/mm^3 (Case-G1, $n = 114$) and $\geq$1,000/mm^3 (Control-G, $n = 302$) at week 2 or 4 during IFN-based therapy. Here there may still be mixed with undesirable samples that should be removed from the case group. Therefore, we designed and carried out the second stage of GWAS (GWAS-2nd) comparing the patients with more severe neutropenia to the control group: in patients with minimum neutrophil counts of <600/mm^3 (Case-G2, $n = 50$) and $\geq$1,000/mm^3 (Control-G, $n = 302$) at week 2 or 4 using the same samples as used in GWAS-1st. Supplementary Fig. 1 shows a genome-wide view of the single-point association data based on allele frequencies in GWAS-1st and GWAS-2nd. No association between SNPs and IFN-induced neutropenia reached a genome-wide level of significance [Bonferroni criterion $P < 8.31 \times 10^{-8}$ (0.05/601,578)]. Therefore, we selected the candidate SNPs principally

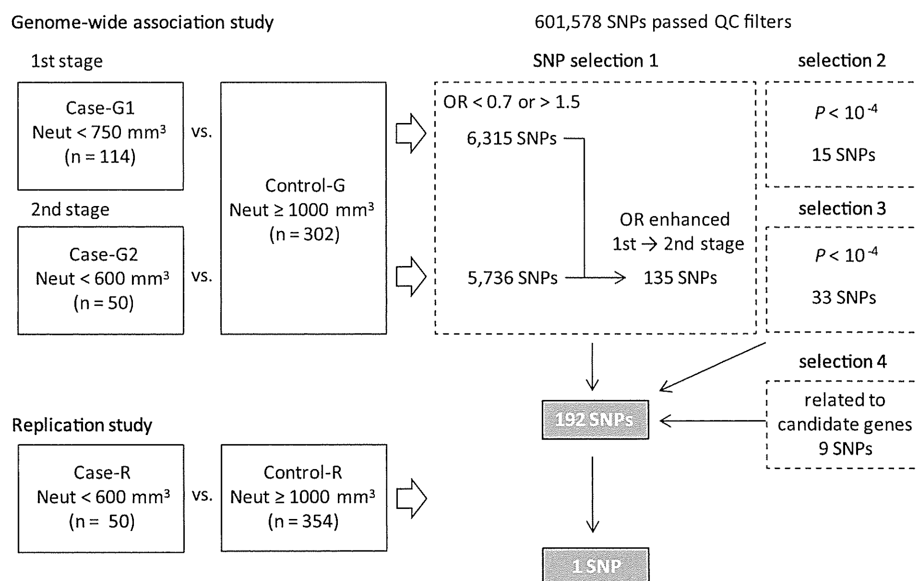

Fig. 1 Outline of the study design. *Neut* neutrophil counts, *SNP* single nucleotide polymorphism, *QC* quality control, *OR* odds ratio

Table 1 Clinical characteristics of patients in GWAS and the replication study

	GWAS			Replication study	
	Case-G1 ($n = 114$)	Case-G2 ($n = 50$)	Control-G ($n = 302$)	Case-R ($n = 50$)	Control-R ($n = 354$)
At baseline					
Gender, male/female	48/66	21/29	170/132	24/26	208/146
Age, years	57.9 (8.7)	57.1 (8.3)	57.2 (11.2)	59.1 (10.2)	56.7 (9.6)
Neutrophil count, /mm^3	1,800 (777)	1,662 (897)	2,750 (984)	1,570 (552)	2,724 (985)
Hemoglobin, g/dL	13.6 (1.3)	13.5 (1.3)	14.2 (1.5)	13.6 (1.6)	14.3 (1.5)
Platelet count, $\times 10^9$/L	141 (42)	132 (46)	164 (54)	140 (47)	162 (60)
ALT, IU/L	82.9 (88.6)	70.4 (53.1)	81.5 (77.9)	87.8 (82.7)	85.2 (71.1)
HCV genotype, 1/2/ND	95/18/1	40/10/0	250/51/1	45/5/0	277/77/0
HCV RNA, log IU/mL	5.9 (0.8)	5.9 (1.0)	6.1 (0.8)	6.1 (0.9)	6.1 (0.8)
Liver fibrosis, F0-2/F3-4/ND	62/22/30	25/10/15	168/70/64	21/6/23	229/87/38
rs8099917, TT/TG + GG/ND	74/39/1	35/15/0	189/109/4	31/17/2	278/70/6
Regimen					
PEG-IFN + RBV/IFN + RBV/PEG-IFN/IFN mono	112/0/0/2	48/0/0/2	277/9/9/7	44/4/2/0	351/0/3/0
At week 4					
Neutrophil count, /mm^3	606 (126)	496 (104)	1,551 (501)	501 (89)	1,533 (484)

Data are expressed as number for categorical data or the mean (standard deviation) for non-categorical data

GWAS genome-wide association study, *ALT* alanine transaminase, *ND* not determined, *PEG-IFN* pegylated interferon, *IFN* mono, interferon monotherapy, *RBV* ribavirin

by comparing between GWAS-1st and GWAS-2nd as follows. There were 6,315 and 5,736 SNPs with odds ratios (ORs) <0.7 or >1.5 at GWAS-1st and GWAS-2nd, respectively. Of these, the ORs of 135 SNPs were more notable at GWAS-2nd than at GWAS-1st. In addition to the 135 SNPs, we selected 15 and 33 SNPs with $P < 10^{-4}$ at GWAS-1st and GWAS-2nd, and added 9 SNPs which are located around the candidate genetic regions identified by the GWAS stage and are non-synonymous or related to diseases in previous reports. Consequently, we carried out the replication analysis focusing on this total of 192 SNPs.

In the subsequent replication analysis, we carried out genotyping of the 192 candidate SNPs in an independent set of 404 Japanese HCV-infected patients with minimum neutrophil counts of <600/mm^3 (Case-R, $n = 50$) and $\geq$1,000/mm^3 (Control-R, $n = 354$) at week 2 or 4 during IFN-based therapy (Table 1; Fig. 1). The results in the replication stage combined with GWAS-2nd are shown in Supplementary Table 1. Several SNPs such as rs11743919 and rs2457840 showed strong associations with low P value, however, the MAF of them were <5 %. In general, low frequent SNPs tend to show unsettled associations, especially in statistical analysis with small number of samples. Therefore, we excluded these SNPs from the final candidates. Consequently, we determined the SNP rs2305482, located in the intron of *PSMD3* gene on chromosome 17, as the most promising candidate, which showed a strong association with IFN-induced neutropenia in the combined results of GWAS-2nd and the replication stage (OR = 2.18; 95 % CI = 1.61–2.96, $P = 3.05 \times 10^{-7}$ in the allele frequency model) (Table 2).

Association of SNPs located in *PSMD3-CSF3* with neutropenia

A previous GWAS showed that rs4794822 located between the *PSMD3* and *CSF3* genes was associated with neutrophil counts in Japanese patients including 14 different disease groups (Okada et al. 2010). As shown in Fig. 2, rs4794822 is in strong linkage disequilibrium (LD) with rs2305482 which we identified in the present study. Thus, the pairwise LD (r^2) in the HapMap JPT: Japanese in Tokyo, Japan, is 0.66. Because the SNP rs4794822 is not included in the Affymetrix Genome-Wide Human SNP Array 6.0, we additionally genotyped it together with three other SNPs (rs9915252, rs3859192 and rs3907022) located in the same LD block around the *PSMD3* gene (Fig. 2). The allele frequency of each SNP was compared between patients with minimum neutrophil counts of <600/mm^3 (Case-G2 + R: Case-G2 plus Case-R, $n = 100$) and $\geq$1,000/mm^3 (Control-G + R: Control-G plus Control-R, $n = 656$) at week 2 or 4 during IFN-based therapy. This showed that, rs4794822 was also strongly associated with neutropenia during IFN-based therapy (OR = 2.24; 95 % CI = 1.63–3.07, $P = 3.63 \times 10^{-7}$ in the allele frequency model) (Table 3).

Table 2 SNP associated with interferon-induced neutropenia

dbSNP rsID	Nearest gene	Risk allele	Allele (1/2)	Stage	Case			Control			OR[a] (95 % CI)	*P* value[b]
					11	12	22	11	12	22		
rs2305482	*PSMD3*	C	C/A	GWAS-1st	23 (20.4)	52 (46.0)	38 (33.6)	26 (8.6)	143 (47.4)	133 (44.0)	1.61 (1.17–2.20)	2.95×10^{-3}
				GWAS-2nd	12 (24.5)	28 (57.1)	9 (18.4)	26 (8.6)	143 (47.4)	133 (44.0)	2.37 (1.54–3.65)	6.47×10^{-5}
				Replication	12 (24.4)	20 (40.8)	17 (34.7)	33 (9.5)	136 (39.1)	179 (51.4)	1.99 (1.30–3.06)	1.46×10^{-3}
				Combined[c]	24 (24.5)	48 (49.0)	26 (26.5)	59 (9.1)	279 (42.9)	312 (48.0)	2.18 (1.61–2.96)	3.05×10^{-7}

Data of allele distribution represent number (%). Data of subjects whose genotypes were not determined were excluded

SNP single nucleotide polymorphism

[a] Odds ratio for the allele frequency model

[b] *P* value by the Chi square test for the allele frequency model

[c] Allele distributions in GWAS-2nd and replication were combined

Predictive factors for IFN-induced neutropenia

The following analyses were carried out for rs2305482 and rs4794822 using the subjects in Case-G2 + R and Control-G + R. Neutrophil counts at baseline correlated with rs2305482 and rs4794822 genotypes (Supplementary Fig. 2), and strongly affected IFN-induced neutropenia as shown by ROC analysis (area under the curve = 0.860) (Supplementary Fig. 3). Furthermore, gender, hemoglobin level, and platelet count at baseline were also significantly associated with IFN-induced neutropenia by univariate analysis (Table 4). Therefore, we analyzed pretreatment predictive factors for IFN-induced neutropenia in logistic regression models that included the following variables: gender, neutrophil count, platelet count, and rs2305482 or rs4794822 genotypes. In addition to neutrophil count, rs2305482 CC was an independent predictive factor for IFN-induced neutropenia (OR = 2.497; 95 % CI = 1.281–4.864, $P = 0.0072$) (Table 5) as was rs4794822 CC (OR = 2.272; 95 % CI = 1.337–3.861, $P = 0.0024$) (Supplementary Table 2).

Impact of *PSMD3-CSF3* SNPs on tolerated drug doses and treatment efficacy

To evaluate the impact of *PSMD3-CSF3* SNPs on doses of drugs given, and on treatment efficacy, we selected 380 HCV genotype 1-infected patients treated with PEG-IFN/RBV for 48 weeks. They were selected as having information available on the doses of PEG-IFN/RBV that they had received (Supplementary Table 3). It was reported that rates of viral clearance were significantly reduced in patients who could not be maintained on at least 80 % of their drug doses for the duration of PEG-IFN/RBV therapy (McHutchison et al. 2002). In reference to this result, we stratified the patients into three groups according to the doses of PEG-IFN or RBV administered, as follows: <60 %, ≥60 to <80 %, ≥80 % of the planned doses for 48 weeks. The proportion of patients in the <60 % group for PEG-IFN was significantly higher in patients possessing rs2305482 CC than in those with AA/AC ($P = 0.005$), whereas there was no association for RBV (Fig. 3). The same results were found in the analysis of rs4794822 (Supplementary Fig. 4). However, the univariate analysis of pretreatment factors associated with SVR showed that there was no association between SVR and rs2305482 or rs4794822 genotypes (Supplementary Table 3).

Candidate SNP-gene association analysis in IFN-induced neutropenia

To investigate whether the SNPs associated with neutropenia affect the expression of nearby genes, we conducted

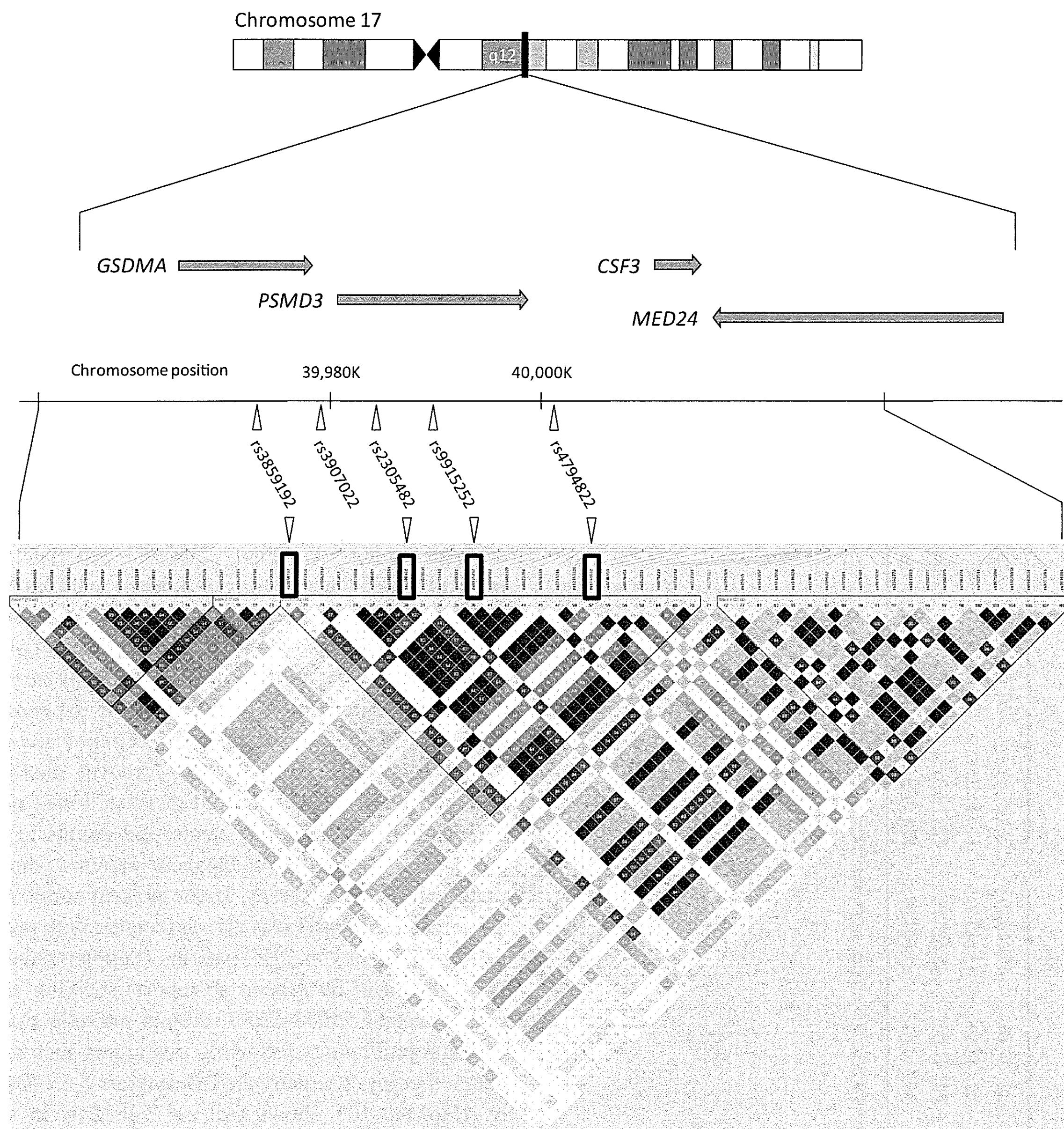

Fig. 2 Position on chromosome and pairwise linkage disequilibrium (r^2) diagrams in the HapMap JPT around the *PSMD3-CSF3* locus

an eQTL analysis. The C allele of rs2305482, a risk for neutropenia, was associated with higher expression levels of *PSMD3* in the populations of LWK: Luhya in Webuye, Kenya (rho = 0.30, $P = 0.006$), and MEX: Mexican ancestry in Los Angeles, California (rho = 0.36, $P = 0.015$) (Supplementary Fig. 5a), whereas it was associated with lower expression levels of *CSF3* in CHB: Han Chinese in Beijing, China, in the probe of ILMN_1655639 (rho = −0.48, $P = 5.5 \times 10^{-6}$) (Supplementary Fig. 5b), and in MEX in that of ILMN_1706852 (rho = −0.33, $P = 0.028$) (Supplementary Fig. 5c).

CSF3 encodes a cytokine, known as G-CSF which is produced by different type of cells such as macrophages, monocytes, stromal cells in the bone marrow, fibroblast, and endothelial cells. The eQTL analysis is based on the whole-genome gene expression variations in lymphoblastoid cell lines derived from HapMap individuals. Therefore, it was still necessary to analyze gene expression in G-CSF producing cells, as well as expression at the protein level. Hence, we measured serum G-CSF levels at baseline and week 2 or 4 (at the time of minimum neutrophil counts) in 127 CHC patients receiving IFN-based therapy. There were no differences in serum G-CSF levels at baseline and the time of minimum neutrophil counts as well as in their changes according to rs2305482 or rs4794822 genotypes (Supplementary Fig. 6a, b). In addition, neutrophil counts

Table 3 Association of SNPs located in *PSMD3-CSF3* with interferon-induced neutropenia

dbSNP rsID	Nearest gene	Risk allele	Allele (1/2)	Case-G2 + R[a] ($n = 100$)			Control-G + R[b] ($n = 656$)			OR[c] (95 % CI)	*P* value[d]
				11	12	22	11	12	22		
rs9915252	*PSMD3*	G	G/C	23 (24.0)	47 (49.0)	26 (27.1)	57 (8.9)	276 (43.3)	304 (47.7)	2.13 (1.57–2.89)	9.64×10^{-7}
rs4794822	*PSMD-CSF3*	C	C/T	42 (42.9)	45 (45.9)	11 (11.2)	130 (21.2)	308 (50.2)	176 (28.7)	2.24 (1.63–3.07)	3.63×10^{-7}
rs3907022	*GSDMA-PSMD*	A	A/G	41 (41.8)	45 (45.9)	12 (12.2)	129 (21.3)	306 (50.6)	170 (28.1)	2.11 (1.54–2.89)	2.31×10^{-6}
rs3859192	*GSDMA*	C	C/T	37 (37.8)	44 (44.9)	17 (17.3)	123 (19.9)	313 (50.7)	181 (29.3)	1.82 (1.34–2.48)	1.04×10^{-4}

Data of allele distribution represent number (%). Data of subjects whose genotypes were not determined were excluded

SNP single nucleotide polymorphism

[a] Case-G2 + R: Case-G2 plus Case-R

[b] Control-G + R: Control-G plus Control-R

[c] Odds ratio for the allele frequency model

[d] *P* value by the Chi square test for the allele frequency model

did not correlate with serum G-CSF levels at baseline and the time of minimum neutrophil counts (Supplementary Fig. 7a), and there was no difference in the changes of serum G-CSF levels from baseline to the time of minimum neutrophil counts between patients with minimum neutrophil counts of $\geq$1,000/mm^3 and <600/mm^3 (Supplementary Fig. 7b).

Discussion

The present GWAS first showed a strong association between genetic variant and IFN-induced neutropenia, namely, with rs2305482 in *PSMD3* on chromosome 17. Although neutrophil counts at baseline were associated with the rs2305482 genotype and the incidence of neutropenia during IFN-based therapy, the logistic regression analysis revealed that the rs2305482 genotype was independently associated with IFN-induced neutropenia.

Intriguingly, the *PSMD3-CSF3* locus was reported to be associated with total white blood cell (WBC) counts based on GWAS of populations with European ancestry (Crosslin et al. 2012; Soranzo et al. 2009) and in Japanese (Kamatani et al. 2010). These findings were replicated in African Americans (Reiner et al. 2011). Moreover, another GWAS by Okada et al. (2010) showed that rs4794822 in *PSMD3-CSF3* was associated with neutrophil counts in 14 different groups of diseases in Japanese patients who were not undergoing chemotherapy. In the present study, rs4794822 as well as rs2305482 was also associated with pretreatment neutrophil counts in CHC patients (Supplementary Fig. 2). However, there have been no reports showing an association between *PSMD3-CSF3* variants and reduction of WBC or neutrophil counts following treatments such as IFN and chemotherapy. The pairwise LD diagram for *PSMD3-CSF3* by HapMap JPT shows that rs4794822 is in strong LD with rs2305482, which we identified here (Fig. 2). In the present study, both rs2305482 and rs4794822 were associated with IFN-induced neutropenia. Collectively, previous reports together with our results imply that the *PSMD3-CSF3* locus is associated with neutropenia in CHC patients under IFN-based therapy as well as with neutrophil counts in healthy individuals and patients without bone marrow suppressive therapy.

In further clinical investigation, the rs2305482 and rs4794822 genotypes were associated with the doses of PEG-IFN that could be given to HCV genotype 1-infected patients treated with PEG-IFN/RBV (Fig. 3; Supplementary Fig. 4). Unfortunately, we could not collect the detailed information about the reason for the reduction of PEG-IFN in this group. However, we highly suppose that these SNPs affected the doses of PEG-IFN through neutropenia in some cases, since neutropenia is one of the major

Table 4 Univariate analysis of pretreatment factors associated with interferon-induced neutropenia

	Case-G2 + R[a] ($n = 100$)	Control-G + R[b] ($n = 656$)	*P* value[c]
Gender, male/female	45/55	378/278	0.018
Age, years	58.1 (9.3)	56.9 (10.4)	0.262
Neutrophil count, /mm^3	1,614 (735)	2,742 (979)	<0.001
Hemoglobin, g/dL	13.5 (1.5)	14.2 (1.5)	<0.001
Platelet count, $\times 10^9$/L	136 (46)	163 (57)	<0.001
ALT, IU/L	79.1 (69.7)	83.5 (74.3)	0.574
HCV RNA, log IU/ml	6.0 (0.9)	6.1 (0.8)	0.164
Liver fibrosis, F0-2/F3-4/ND	46/16/38	397/157/102	0.674
rs2305482, AA + AC/CC/ND	74/24/2	591/59/6	<0.001
rs4794822, TT + TC/CC/ND	56/42/2	484/130/42	<0.001

Data are expressed as number for categorical data or the mean (standard deviation) for non-categorical data

ALT alanine transaminase, *ND* not determined

[a] Case-G2 + R: Case-G2 plus Case-R

[b] Control-G + R: Control-G plus Control-R

[c] Categorical variables were compared between groups by the Chi square test and non-categorical variables by the Student's *t* test

Table 5 Logistic regression analysis of pretreatment factors associated with interferon-induced neutropenia

	OR (95 % CI)	*P* value
Gender, female	1.229 (0.734–2.059)	0.4331
Neutrophil count, /mm^3	0.998 (0.997–0.998)	<0.0001
Platelet count, $\times 10^9$/L	1.005 (0.953–1.059)	0.8604
rs2305482, CC	2.497 (1.281–4.864)	0.0072

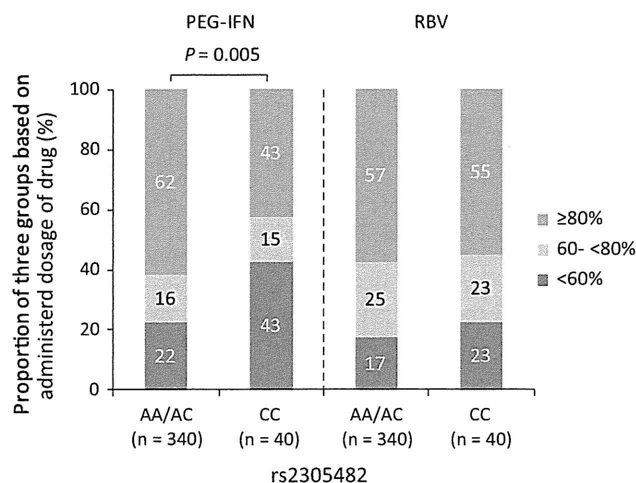

Fig. 3 Administered doses of PEG-IFN and RBV according to rs2305482 genotypes. The patients were stratified into three groups according to the doses of PEG-IFN or RBV administered, as follows: <60 %, ≥60 to <80 %, ≥80 % of the planned doses for 48 weeks. The proportion of patients receiving <60 % of the PEG-IFN doses was significantly higher in patients with rs2305482 CC than in those with AA/AC ($P = 0.005$, by the Chi square test). *PEG-IFN* pegylated interferon, *RBV* ribavirin

reasons for the dose reduction of PEG-IFN in PEG-IFN/RBV therapy. While, there were no associations between SVR and rs2305482 or rs4794822 genotypes (Supplementary Table 3).

PSMD3 encodes the proteasome 26S subunit, non-ATPase 3, a member of the 26S proteasome family, and is involved in the control of cell cycle transition via the ubiquitin–proteasome pathway (Bailly and Reed 1999). *CSF3* encodes G-CSF, which controls the production, differentiation, and function of granulocytes (Nagata et al. 1986). Recombinant G-CSF is widely used to treat patients with severe neutropenia during chemotherapy. Therefore, we hypothesize that *PSMD3*-*CSF3* variants may influence neutrophil counts through affecting the process of endogenous G-CSF synthesis during IFN-based therapy or other bone marrow suppressive therapies. However, eQTL analysis by Okada et al. (2010) showed that rs4794822 was significantly associated with the expression level of *PSMD3*, rather than that of *CSF3* in the JPT and CHB populations. Our eQTL analysis showed that the risk allele for neutropenia at rs2305482 correlated with higher expression levels of *PSMD3* in LWK and MEX populations (Supplementary Fig. 5a), whereas with lower expression levels of *CSF3* in MEX and especially in CHB populations (Supplementary Fig. 5b, c). However, these results were not replicated in the other probe of *CSF3*. Additionally, we analyzed serum G-CSF levels in CHC patients receiving IFN-based therapy. Although serum G-CSF levels were thought to be increased in response to neutropenia regardless of rs2305482 and rs4794822 genotypes, there was no evidence that they were lower in patients with a risk allele of these SNPs at baseline and during the neutropenic period (Supplementary Fig. 6). Moreover, neutrophil counts did not correlate with serum

G-CSF levels at baseline and the time of minimum neutrophil counts (Supplementary Fig. 7a). Further functional analyses of these genes and polymorphisms are required to elucidate the reason for the association between *PSMD3-CSF3* and IFN-induced neutropenia as well as neutrophil counts in healthy individuals.

In previous reports, *PLBC4*, *DARC*, *CXCL2*, and *CDK5* loci have also been associated with neutrophil or WBC counts in healthy individuals or patients who were not under chemotherapy (Crosslin et al. 2012; Kamatani et al. 2010; Okada et al. 2010; Reiner et al. 2011). However, there were no associations with these loci discernible in our GWAS.

The important limitation of this study is that the association between rs2305482 and IFN-induced neutropenia was not statistically significant in a genome-wide level. Thompson et al. (2012) also identified no genetic determinants of IFN-induced neutropenia during PEG-IFN/RBV therapy at the level of genome-wide significance by their GWAS. Unlike our study design, they analyzed the association between the reduction of neutrophil counts at week 4 and any SNPs. Indeed, we analyzed the association between the reduction of neutrophil counts at week 2 or 4 and rs2305482 or rs4794822, but there was no significant association. Therefore, further independent replication analyses which are designed in the similar way as our study are desirable.

IFN-free therapies are expected to be useful especially in IFN-resistant patients and may become the standard of care in the near future. However, combination therapies of DAA and IFN will continue to be used for some time. Our findings contribute to our understanding of the genetic factors influencing IFN-induced neutropenia. Furthermore, these genetic variants may be associated with neutropenia during chemotherapies for various malignant diseases as well as IFN-based therapy for CHC. Therefore, genetic testing of these variants might be useful for establishing personalized doses of such therapies to minimize drug-induced adverse events. Additionally, our results might contribute to the elucidation of the mechanism of drug-induced neutropenia.

Acknowledgments We thank Ms. Yasuka Uehara-Shibata, Yuko Ogasawara-Hirano, Yoshimi Ishibashi, Natsumi Baba, Megumi Yamaoka-Sageshima, Takayo Tsuchiura, Yoriko Mawatari (Tokyo University), and Dr. Shintaro Ogawa (Nagoya City University) for technical assistance. This work was supported by the Ministry of Health, Labor, and Welfare of Japan (H25-kanen-ippan-005) to Yasuhito Tanka and Katsushi Tokunaga, KAKENHI (22133008) Grant-in-Aid for Scientific Research on Innovative Areas to Katsushi Tokunaga, and KAKENHI (24790728) Grant-in-Aid from the Ministry of Education, Culture, Sports, Science of Japan for Young Scientists (B) to Nao Nishida.

Conflict of interest The following authors are currently conducting research sponsored by the companies: Yasuhito Tanaka, Keisuke Hino, and Yoshito Itoh by Merck Sharp & Dohme, Corp., Chugai Pharmaceutical Co., Ltd., and Bristol-Myers Squibb; Nobuyuki Enomoto, Shuhei Nishiguchi, and Eiji Tanaka by Merck Sharp & Dohme, Corp. and Chugai Pharmaceutical Co., Ltd.; Naoya Sakamoto by Chugai Pharmaceutical Co., Ltd, Bristol-Myers Squibb, Merck Sharp & Dohme, Corp., and Otsuka Pharmaceutical Co., Ltd.; Hiroshi Yatsuhashi by Chugai Pharmaceutical Co., Ltd.; Akihiro Tamori by Merck Sharp & Dohme, Corp.; Satoshi Mochida by Merck Sharp & Dohme, Corp., Chugai Pharmaceutical Co., Ltd., Bristol-Myers Squibb, and Toray Medical Co., Ltd. The other authors have no conflict of interest.

Compliance with ethical standards All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Bailly E, Reed SI (1999) Functional characterization of rpn3 uncovers a distinct 19S proteasomal subunit requirement for ubiquitin-dependent proteolysis of cell cycle regulatory proteins in budding yeast. Mol Cell Biol 19:6872–6890

Crosslin DR, McDavid A, Weston N, Nelson SC, Zheng X, Hart E, de Andrade M, Kullo IJ, McCarty CA, Doheny KF, Pugh E, Kho A, Hayes MG, Pretel S, Saip A, Ritchie MD, Crawford DC, Crane PK, Newton K, Li R, Mirel DB, Crenshaw A, Larson EB, Carlson CS, Jarvik GP (2012) Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. Hum Genet 131:639–652. doi:10.1007/s00439-011-1103-9

Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, Little LD, Qiu P, Bertelsen AH, Watson M, Warner A, Muir AJ, Brass C, Albrecht J, Sulkowski M, McHutchison JG, Goldstein DB (2010) ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature 464:405–408. doi:10.1038/nature08825

Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461:399–401. doi:10.1038/nature08309

George SL, Bacon BR, Brunt EM, Mihindukulasuriya KL, Hoffmann J, Di Bisceglie AM (2009) Clinical, virologic, histologic, and biochemical outcomes after successful HCV therapy: a 5-year follow-up of 150 patients. Hepatology 49:729–738. doi:10.1002/hep.22694

Jacobson IM, McHutchison JG, Dusheiko G, Di Bisceglie AM, Reddy KR, Bzowej NH, Marcellin P, Muir AJ, Ferenci P, Flisiak R, George J, Rizzetto M, Shouval D, Sola R, Terg RA, Yoshida EM, Adda N, Bengtsson L, Sankoh AJ, Kieffer TL, George S, Kauffman RS, Zeuzem S (2011) Telaprevir for previously untreated chronic hepatitis C virus infection. N Engl J Med 364:2405–2416. doi:10.1056/NEJMoa1012912

Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, Nakamura Y, Kamatani N (2010) Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat Genet 42:210–215. doi:10.1038/ng.531

Kurosaki M, Tanaka Y, Tanaka K, Suzuki Y, Hoshioka Y, Tamaki N, Kato T, Yasui Y, Hosokawa T, Ueda K, Tsuchiya K, Kuzuya T,

Nakanishi H, Itakura J, Takahashi Y, Asahina Y, Matsuura K, Sugauchi F, Enomoto N, Nishida N, Tokunaga K, Mizokami M, Izumi N (2011) Relationship between polymorphisms of the inosine triphosphatase gene and anaemia or outcome after treatment with pegylated interferon and ribavirin. Antivir Ther 16:685–694. doi:10.3851/IMP1796

Matsuura K, Tanaka Y, Watanabe T, Fujiwara K, Orito E, Kurosaki M, Izumi N, Sakamoto N, Enomoto N, Yatsuhashi H, Kusakabe A, Shinkai N, Nojiri S, Joh T, Mizokami M (2014) ITPA genetic variants influence efficacy of PEG-IFN/RBV therapy in older patients infected with HCV genotype 1 and favourable IL28B type. J Viral Hepat 21:466–474. doi:10.1111/jvh.12171

McHutchison JG, Manns M, Patel K, Poynard T, Lindsay KL, Trepo C, Dienstag J, Lee WM, Mak C, Garaud JJ, Albrecht JK (2002) Adherence to combination therapy enhances sustained response in genotype-1-infected patients with chronic hepatitis C. Gastroenterology 123:1061–1069

Nagata S, Tsuchiya M, Asano S, Kaziro Y, Yamazaki T, Yamamoto O, Hirata Y, Kubota N, Oheda M, Nomura H et al (1986) Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. Nature 319:415–418. doi:10.1038/319415a0

Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 364:78–85. doi:10.1016/j.ab.2007.02.005

Ochi H, Maekawa T, Abe H, Hayashida Y, Nakano R, Kubo M, Tsunoda T, Hayes CN, Kumada H, Nakamura Y, Chayama K (2010) ITPA polymorphism affects ribavirin-induced anemia and outcomes of therapy—a genome-wide study of Japanese HCV virus patients. Gastroenterology 139:1190–1197. doi:10.1053/j.gastro.2010.06.071

Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, Daigo Y, Yamamoto K, Kubo M, Nakamura Y, Kamatani N (2010) Common variations in PSMD3-CSF3 and PLCB4 are associated with neutrophil count. Hum Mol Genet 19:2079–2085. doi:10.1093/hmg/ddq080

Poordad F, Bronowicki JP, Gordon SC, Zeuzem S, Jacobson IM, Sulkowski MS, Poynard T, Morgan TR, Molony C, Pedicone LD, Sings HL, Burroughs MH, Sniukiene V, Boparai N, Goteti VS, Brass CA, Albrecht JK, Bacon BR (2012) Factors that predict response of patients with hepatitis C virus infection to boceprevir. Gastroenterology 143(608–18):e1–e5. doi:10.1053/j.gastro.2012.05.011

Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, Dean E, Arepalli S, Britton A, Chen Z, Couper D, Curb JD, Eaton CB, Fornage M, Grant SF, Harris TB, Hernandez D, Kamatini N, Keating BJ, Kubo M, LaCroix A, Lange LA, Liu S, Lohman K, Meng Y, Mohler ER 3rd, Musani S, Nakamura Y, O'Donnell CJ, Okada Y, Palmer CD, Papanicolaou GJ, Patel KV, Singleton AB, Takahashi A, Tang H, Taylor HA Jr, Taylor K, Thomson C, Yanek LR, Yang L, Ziv E, Zonderman AB, Folsom AR, Evans MK, Liu Y, Becker DM, Snively BM, Wilson JG (2011) Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). PLoS Genet 7:e1002108. doi:10.1371/journal.pgen.1002108

Sakamoto N, Tanaka Y, Nakagawa M, Yatsuhashi H, Nishiguchi S, Enomoto N, Azuma S, Nishimura-Sakurai Y, Kakinuma S, Nishida N, Tokunaga K, Honda M, Ito K, Mizokami M, Watanabe M (2010) ITPA gene variant protects against anemia induced by pegylated interferon-alpha and ribavirin therapy for Japanese patients with chronic hepatitis C. Hepatol Res 40:1063–1071. doi:10.1111/j.1872-034X.2010.00741.x

Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, Salo P, Voight BF, Burns P, Laskowski RA, Xue Y, Menzel S, Altshuler D, Bradley JR, Bumpstead S, Burnett MS, Devaney J, Doring A, Elosua R, Epstein SE, Erber W, Falchi M, Garner SF, Ghori MJ, Goodall AH, Gwilliam R, Hakonarson HH, Hall AS, Hammond N, Hengstenberg C, Illig T, Konig IR, Knouff CW, McPherson R, Melander O, Mooser V, Nauck M, Nieminen MS, O'Donnell CJ, Peltonen L, Potter SC, Prokisch H, Rader DJ, Rice CM, Roberts R, Salomaa V, Sambrook J, Schreiber S, Schunkert H, Schwartz SM, Serbanovic-Canic J, Sinisalo J, Siscovick DS, Stark K, Surakka I, Stephens J, Thompson JR, Volker U, Volzke H, Watkins NA, Wells GA, Wichmann HE, Van Heel DA, Tyler-Smith C, Thein SL, Kathiresan S, Perola M, Reilly MP, Stewart AF, Erdmann J, Samani NJ, Meisinger C, Greinacher A, Deloukas P, Ouwehand WH, Gieger C (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet 41:1182–1190. doi:10.1038/ng.467

Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E, Riordan S, Sheridan D, Smedile A, Fragomeli V, Muller T, Bahlo M, Stewart GJ, Booth DR, George J (2009) IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 41:1100–1104. doi:10.1038/ng.447

Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41:1105–1109. doi:10.1038/ng.449

Tanaka Y, Kurosaki M, Nishida N, Sugiyama M, Matsuura K, Sakamoto N, Enomoto N, Yatsuhashi H, Nishiguchi S, Hino K, Hige S, Itoh Y, Tanaka E, Mochida S, Honda M, Hiasa Y, Koike A, Sugauchi F, Kaneko S, Izumi N, Tokunaga K, Mizokami M (2011) Genome-wide association study identified ITPA/DDRGK1 variants reflecting thrombocytopenia in pegylated interferon and ribavirin therapy for chronic hepatitis C. Hum Mol Genet 20:3507–3516. doi:10.1093/hmg/ddr249

Thompson AJ, Clark PJ, Singh A, Ge D, Fellay J, Zhu M, Zhu Q, Urban TJ, Patel K, Tillmann HL, Naggie S, Afdhal NH, Jacobson IM, Esteban R, Poordad F, Lawitz EJ, McCone J, Shiffman ML, Galler GW, King JW, Kwo PY, Shianna KV, Noviello S, Pedicone LD, Brass CA, Albrecht JK, Sulkowski MS, Goldstein DB, McHutchison JG, Muir AJ (2012) Genome-wide association study of interferon-related cytopenia in chronic hepatitis C patients. J Hepatol 56:313–319. doi:10.1016/j.jhep.2011.04.021

Yang TP, Beazley C, Montgomery SB, Dimas AS, Gutierrez-Arcelus M, Stranger BE, Deloukas P, Dermitzakis ET (2010) Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. Bioinformatics 26:2474–2476. doi:10.1093/bioinformatics/btq452

Yoshida H, Tateishi R, Arakawa Y, Sata M, Fujiyama S, Nishiguchi S, Ishibashi H, Yamada G, Yokosuka O, Shiratori Y, Omata M (2004) Benefit of interferon therapy in hepatocellular carcinoma prevention for individual patients with chronic hepatitis C. Gut 53:425–430

Zeuzem S, Andreone P, Pol S, Lawitz E, Diago M, Roberts S, Focaccia R, Younossi Z, Foster GR, Horban A, Ferenci P, Nevens F, Mullhaupt B, Pockros P, Terg R, Shouval D, van Hoek B, Weiland O, Van Heeswijk R, De Meyer S, Luo D, Boogaerts G, Polo R, Picchio G, Beumont M (2011) Telaprevir for retreatment of HCV infection. N Engl J Med 364:2417–2428. doi:10.1056/NEJMoa1013086

JVI
Journals.ASM.org

Hallmarks of Hepatitis C Virus in Equine Hepacivirus

Tomohisa Tanaka,[a] Hirotake Kasai,[a] Atsuya Yamashita,[a] Kaori Okuyama-Dobashi,[a] Jun Yasumoto,[a] Shinya Maekawa,[b] Nobuyuki Enomoto,[b] Toru Okamoto,[c] Yoshiharu Matsuura,[c] Masami Morimatsu,[d] Noboru Manabe,[e] Kazuhiko Ochiai,[f] Kazuto Yamashita,[g] Kohji Moriishi[a]

Department of Microbiology, Faculty of Medicine, University of Yamanashi, Yamanashi, Japan[a]; First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, Yamanashi, Japan[b]; Department of Molecular Virology, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan[c]; Laboratory of Laboratory Animal Science and Medicine, Department of Disease Control, Graduate School of Veterinary Medicine, Hokkaido University, Sapporo, Japan[d]; Animal Resource Science Center, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Kasama, Japan[e]; Department of Basic Science, School of Veterinary Nursing and Technology, Faculty of Veterinary Science, Nippon Veterinary and Life Science University, Tokyo, Japan[f]; Department of Small Animal Clinical Sciences, School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu, Hokkaido, Japan[g]

ABSTRACT

Equine hepacivirus (EHcV) has been identified as a closely related homologue of hepatitis C virus (HCV) in the United States, the United Kingdom, and Germany, but not in Asian countries. In this study, we genetically and serologically screened 31 serum samples obtained from Japanese-born domestic horses for EHcV infection and subsequently identified 11 PCR-positive and 7 seropositive serum samples. We determined the full sequence of the EHcV genome, including the 3′ untranslated region (UTR), which had previously not been completely revealed. The polyprotein of a Japanese EHcV strain showed approximately 95% homology to those of the reported strains. HCV-like *cis*-acting RNA elements, including the stem-loop structures of the 3′ UTR and kissing-loop interaction were deduced from regions around both UTRs of the EHcV genome. A comparison of the EHcV and HCV core proteins revealed that Ile^{190} and Phe^{191} of the EHcV core protein could be important for cleavage of the core protein by signal peptide peptidase (SPP) and were replaced with Ala and Leu, respectively, which inhibited intramembrane cleavage of the EHcV core protein. The loss-of-function mutant of SPP abrogated intramembrane cleavage of the EHcV core protein and bound EHcV core protein, suggesting that the EHcV core protein may be cleaved by SPP to become a mature form. The wild-type EHcV core protein, but not the SPP-resistant mutant, was localized on lipid droplets and partially on the lipid raft-like membrane in a manner similar to that of the HCV core protein. These results suggest that EHcV may conserve the genetic and biological properties of HCV.

IMPORTANCE

EHcV, which shows the highest amino acid or nucleotide homology to HCV among hepaciviruses, was previously reported to infect horses from Western, but not Asian, countries. We herein report EHcV infection in Japanese-born horses. In this study, HCV-like RNA secondary structures around both UTRs were predicted by determining the whole-genome sequence of EHcV. Our results also suggest that the EHcV core protein is cleaved by SPP to become a mature form and then is localized on lipid droplets and partially on lipid raft-like membranes in a manner similar to that of the HCV core protein. Hence, EHcV was identified as a closely related homologue of HCV based on its genetic structure as well as its biological properties. A clearer understanding of the epidemiology, genetic structure, and infection mechanism of EHcV will assist in elucidating the evolution of hepaciviruses as well as the development of surrogate models for the study of HCV.

The *Flaviviridae* family is composed of four genera: *Flavivirus*, *Pestivirus*, *Pegivirus*, and *Hepacivirus*. *Flaviviridae* family viruses are enveloped and contain a single-stranded, positive-sense RNA genome, which encodes a single large precursor polyprotein composed of approximately 2,800 to 3,000 amino acids. The genus *Hepacivirus* had included only two species, hepatitis C virus (HCV) and GB virus B (GBV-B), until 2010. GBV-B was isolated from serum samples obtained from laboratory tamarins by 11 passages of serum obtained from a human patient with idiopathic hepatitis (1). Although GBV-B experimentally infects tamarins and common marmosets, but not chimpanzees, *in vivo* (2, 3), the natural host of GBV-B has not yet been clarified. Several hepacivirus species were recently detected in dogs, horses, bats, and rodents and tentatively designated nonprimate hepaciviruses (NPHVs). Bat hepaciviruses have been isolated from some species of bats in Kenya (4), while rodent hepaciviruses have been isolated from several species of rodents in Germany, the Netherlands, South Africa, and Namibia (5, 6). GBV-B is phylogenetically more similar to rodent hepacivirus than to HCV (5). Several strains of equine hepacivirus (EHcV) have been isolated from domestic horses in the United States, the United Kingdom, and Germany (5, 7, 8). The canine hepacivirus was isolated from dogs in the United States (9) but has not yet been genetically or serologically detected in any dogs other than those from the first report (5, 7, 8). The polypeptides of canine hepacivirus show approximately 95% amino acid homology to those of the EHcV strains, suggesting that canine hepacivirus may belong to the same species as EHcV and

Received 8 August 2014 Accepted 2 September 2014
Published ahead of print 10 September 2014
Editor: T. S. Dermody
Address correspondence to Kohji Moriishi, kmoriishi@yamanashi.ac.jp.
Copyright © 2014, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JVI.02280-14

that infections may be rare in dogs (5, 7, 8, 10). Recent phylogenetic analyses identified EHcV as the most closely related viral homologue of HCV among the reported NPHV strains; however, epidemiological and virological information on EHcV is limited. The open reading frames of EHcV strains show approximately 95% homology to one another, suggesting that previously reported EHcV strains may be classified into one species. Several genome sequences of rodent hepacivirus have already been completely determined (5). The 3′ untranslated region (UTR) of HCV was found to include three stem-loop (SL) structures, while variable stem-loop structures were found in that of rodent hepacivirus and GBV-B (5). However, the nucleotide sequence of the EHcV 3′ UTR has not yet been determined completely because the adenine-rich [(A)-rich] sequence downstream of the stop codon in the EHcV genome interrupts an ordinary 3′-rapid amplification of cDNA ends (RACE) reaction (8). The RNA secondary structure of the hepacivirus 3′ UTR may indicate species specificity (5).

On the basis of amino acid similarities among the polyproteins of NPHVs and HCV, the N-terminal one-fourth of the NPHV polyprotein has been predicted to be cleaved by signal peptidase into mature structural proteins and a viroporin (core, E1, E2, and p7), while the C-terminal three-fourths has been predicted to be cleaved by viral proteases into maturated nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (6). Core, E1, and E2 have been predicted to form viral particles with host lipids, although it remains unclear whether p7 is incorporated into a viral particle. Signal peptide peptidase (SPP) was shown to further cleave the C-terminal transmembrane region of HCV and GBV-B core protein after signal peptidase-dependent cleavage (11, 12). However, whether SPP cleaves the C-terminal transmembrane region of the NPHV core protein remains unknown.

The mature core proteins of HCV and GBV-B are localized mainly on lipid droplets (LDs) (13, 14). The core proteins of dengue virus are also localized on LDs but are not cleaved by SPP (15), suggesting that localization of the core protein on LDs may be one of the common characteristics of the *Flaviviridae* family. The HCV core protein is known to be partially localized in the detergent-resistant membrane (DRM), which originates from lipid raft-like membranes (16, 17). The DRM is composed of cholesterol and sphingolipids, which are included in the replication compartment known as the membranous web (18, 19). Therefore, LDs and DRM are considered to be the intracellular compartments for the replication and viral assembly of HCV, but it is currently unknown whether NPHV core proteins are localized on LDs and DRM.

Epidemiological information on EHcV is still limited. The results of the present study demonstrated that Japanese-born domestic horses were infected with EHcV, which showed high homology to the reported strains on the basis of its nucleotide and amino acid sequences. We predicted the RNA secondary structures around the 5′ and 3′ UTRs of the EHcV genome and analyzed the biological properties of the EHcV core protein in relation to the HCV core protein.

MATERIALS AND METHODS

Samples. Serum samples 1 to 13 were collected from Japanese-born domestic horses raised on one farm, farm A, located in Hokkaido, Japan, while groups of serum samples numbered 14 to 18 and 19 to 31 were from horses on farms B and C, respectively, located in Tokyo, Japan (Fig. 1). The distance between Hokkaido and Tokyo is about 1,000 km. All sample collections conformed to guidelines for the care and use of laboratory animals (Yamanashi University) and were approved by the Institutional Committee of Laboratory Animal Experimentation (Yamanashi University). All samples were divided into small aliquots and stored at −80°C until nucleic acid extraction.

RT-PCR. Total RNAs were prepared from horse sera using a Qiagen viral RNA extraction kit (Qiagen, Valencia, CA). RNAs were converted to cDNA using a PrimeScript reverse transcription-PCR (RT-PCR) kit (TaKaRa, Shiga, Japan) with random primers. The viral gene was amplified by PCR using PuReTaq Ready-To-Go PCR beads (GE Healthcare, Piscataway, NJ) with three pairs of primers: NPHV-F1 (5′-TGTCACCT ACTATCGGGG-3′) and NPHV-R1 (5′-TCAAGCCTATACAGCAAAG G-3′), NPHV-F2 (5′-ATCATTTGTGATGAGTGCC-3′) and NPHV-R2 (5′-CATAAGGGCGTCCGTGGC-3′), and NPHV-F3 (5′-GTGGTCGCC ACGGATGCC-3′) and NPHV-R3 (5′-ACCCTATGAAGACGCTCTCC-3′). PCR was carried out as follows: one cycle at 92°C for 5 min; 35 repeats of one cycle at 94°C for 0.5 min, 58°C for 0.5 min, and 72°C for 0.5 min, in that order; and one cycle at 72°C for 1 min followed by holding at 4°C. The PCR products were electrophoresed on 1.5% agarose gels, stained with ethidium bromide, and visualized using the BioDoc-It imaging system (UVP, Upland, CA).

Determination of the EHcV genomic sequence. The viral genome of EHcV was segmentally amplified by PCR using the primers listed in Table 1. The PCR products were cloned into T vectors prepared from pBluescript II SK(−) (20). The DNA sequences of the PCR products were determined using an ABI Prism BigDye Terminator version 1.1 cycle sequencing kit and an ABI Prism 310 genetic analyzer (Life Technologies, Tokyo, Japan). More than three colonies were picked up among the transformants of *Escherichia coli* with regard to the accuracy of the sequence. The nucleotide sequences of the PCR products were determined in forward and reverse directions. The junction of two adjacent PCR products was confirmed by PCR using primers that overlapped two close regions. The 5′-terminal sequence upstream of the open reading frame was determined with a 5′-RACE core set (TaKaRa) using the 5′ phospholyrated RT primer for the NPHV 5′ UTR (5′-CATCCTATCAGACCG-3′). The 3′-terminal region downstream of the (A)-rich region was determined by the 3′-RACE method (21, 22), modified as follows: Total RNAs were prepared from horse serum using TRIzol LS reagent (Invitrogen, Carlsbad, CA) with 40 μg of glycogen (Nacalai Tesque, Kyoto, Japan). The poly(U) tail was added to the 3′ end of the RNA preparation using *Escherichia coli* poly(U) polymerase (New England BioLabs, Ipswich, MA) and was incubated for 45 min at 37°C. The resulting preparation was reverse transcribed by the SuperScript First-Strand Synthesis system (Life Technologies) using an oligo(dA) adapter primer (5′-TTGCGAGCACAGAATTAATACG ACTCACAAAAAAAAAAAAVN-3′). The sequence of each region was determined by sequencing more than 3 clones. The primers for PCR amplification and the RACE methods are listed in Table 1. The whole sequence of the EHcV strain isolated from serum sample 3 (GenBank accession number AB863589) was determined by the method described above. The EHcV strain was designated JPN3/JAPAN/2013 in this study. The partial NS5B-coding regions and 3′ UTRs were amplified from serum samples 5 and 1. The nucleotide sequences of samples 5 and 1 (GenBank accession numbers AB921150 and AB921151, respectively) were determined by the method described above. The neighbor-joining trees of the nucleotide sequences from the NPHV, HCV, and GBV-B strains were predicted by the method of Saitou et al (23). Trees were constructed by the maximum composite likelihood method calculated by using the program MEGA5 (24) (see Fig. 3). The secondary protein structures were predicted by the method of Garnier et al. (25) (see Fig. 6). Hydrophobicity plots of the EHcV and HCV core proteins were prepared by the method of Kyte and Doolittle (26) and drawn using the software Genetyx (Nihon Genetyx, Tokyo, Japan) (see Fig. 5).

Quantification of viral genomic RNAs in horse sera. Total RNA was prepared from equine serum using a Qiagen viral RNA extraction kit and was then reverse transcribed into cDNA by using a PrimeScript RT-PCR

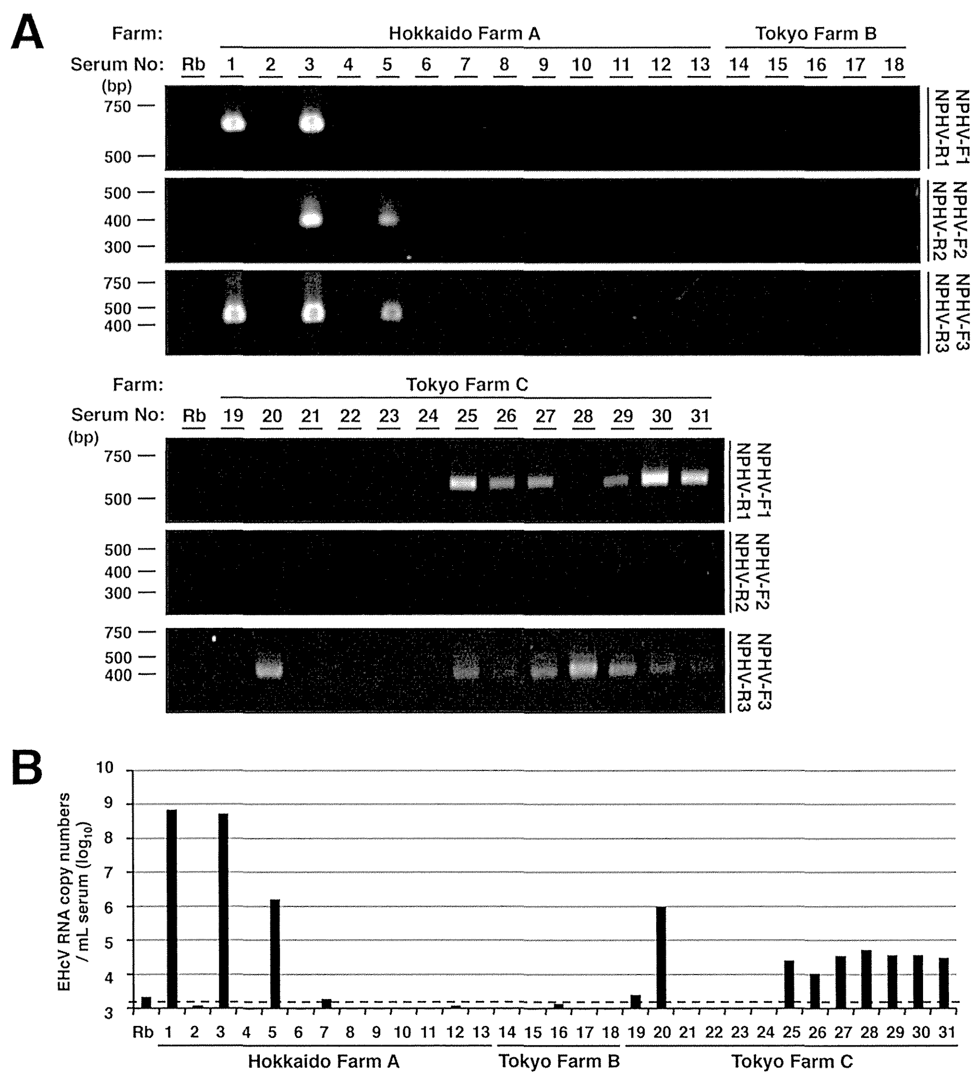

FIG 1 Detection and genetic analyses of NPHV genomic RNA in sera of Japanese domestic horses. (A) Total RNAs extracted from 31 equine sera and normal rabbit serum (Rb) as a negative control were subjected to RT-PCR analysis. Hokkaido Farm A, Tokyo Farm B, and Tokyo Farm C indicate the farms where the individual horses were reared. Three sets of primers, NPHV-F1 and NPHV-R1, NPHV-F2 and NPHV-R2, and NPHV-F3 and NPHV-R3, were used to amplify NPHV-specific gene regions. The PCR products were electrophoresed and stained with ethidium bromide. (B) Total RNAs were isolated from sera, reverse-transcribed, and estimated as a copy number per ml. Normal rabbit serum was used as a negative control. The dashed line indicates the cutoff level.

kit with random primers. The amount of targeted viral RNA was estimated using SYBR GreenER qPCR SuperMix (Life Technologies) and the ABI StepOnePlus real-time PCR system (Life Technologies). The region encoding NS3 was targeted with the primer pair NPHV-F3 (5′-GTGGTC GCCACGGATGCC-3′) and NPHV-R3 (5′-ACCCTATGAAGACGCTC TCC-3′). Total RNAs extracted from conventional rabbit serum were used as a negative control to determine the analytical threshold line. The *in vitro*-transcribed RNA of EHcV was utilized for the standard curve.

Prediction of RNA secondary structures. The 5′-UTR sequences of EHcV strains were aligned with the MUSCLE program and subjected to a manual search for covariant nucleotide substitutions. The RNA folding structure upstream of domain III in the 5′ UTR was predicted using the Mfold web server (27) with conventional phylogenetic conservation analysis due to the lack of sufficient homology to the 5′ UTR sequences of HCV strains. The NS5B-coding regions and 3′ UTRs of EHcV strains were aligned with the program MUSCLE. Conserved secondary structures were predicted as described above. The secondary structures of the 3′ UTR in EHcV were predicted by the Mfold web server without confirming phylogenetic data because of the absence of additional available sequences of the EHcV 3′ UTR and the lack of sufficient homology to the HCV X-tail sequences.

Plasmids. The PCR product encoding the EHcV core protein was amplified from serum sample 3 and was then cloned into the BamHI and XhoI sites of pcDNA3.1-Flag/HA, which encodes the FLAG and hemagglutinin (HA) epitope tags, as reported previously (28). Ala[204] was replaced with Lys to prevent signal peptidase-dependent cleavage. The translated EHcV core protein was added to the FLAG and HA epitope tags at the N and C termini (EHcVc), respectively. A point mutation was generated using a KOD mutagenesis kit (Toyobo, Osaka, Japan). The PCR products encoding EHcVc or the mutant in which Ile[190] and Phe[191] were replaced with Ala and Leu (EHcVc-mt), respectively, were introduced into the AflII and EcoRV sites of pCAGGS using an In-Fusion HD cloning kit (TaKaRa). The introduced fragments of all plasmids were confirmed by sequencing using an ABI Prism 310 genetic analyzer (Applied Biosystems). The plasmid encoding the N-terminally FLAG-tagged and C-terminally HA-tagged HCV core protein (HCVc) and the mutant in which

Downloaded from http://jvi.asm.org/ on February 15, 2015 by guest

TABLE 1 List of PCR primers used in this study

Primer	Nucleotide position	Genome location	F or R[a]	Sequence (5′→3′)
Primer for cloning of the NPHV genome	92–111	5′ UTR	F	ATGTGTCACTCCCCCTATGG
	367–386	5′ UTR	R	CTATGGTCTACGAGACCGGC
	268–285	5′ UTR	F	AGCCGAAATTTGGGCGTG
	1207–1224	E1	R	AAACAGAAGCCATAGCGG
	1116–1132	E1	F	AGTGCTTGTTGGGTGCC
	1697–1713	E2	R	GTCCTTTGCACTTCGGG
	1605–1623	E2	F	ACTGTTAAGCAGATGTGGG
	2103–2121	E2	R	CACAGAGTTGGTAAGTAGC
	2007–2023	E2	F	AAGCAGTGTGGTGCTCC
	2526–2545	E2	R	AAACAGAACCAGAGAATTGC
	2375–2392	E2	F	CCCTGCCTTCACTACTGG
	2898–2913	NS2	R	CGAGATAGCGCCAAGC
	2847–2867	NS2	F	TTTATGCTAGTAAAGTGGTGG
	3396–3415	NS2	R	GGTGATAAAAGTCTCCATCC
	3318–3334	NS2	F	ATCCTCCATGGCTTGCC
	3819–3835	NS3	R	GGGCCACCTGAACTACC
	3732–3750	NS3	F	ACCAGGACGGGTCAGGTCG
	4254–4270	NS3	R	ATAATGTCATAAGCACC
	4177–4195	NS3	F	CTAGTTGCAAGACAACGGG
	4682–4700	NS3	R	AGTGTTGCAGTCAGTGACG
	4574–4591	NS3	F	TGTCACCTACTATCGGGG
	5199–5218	NS3	R	TCAAGCCTATACAGCAAAGG
	4574–4591	NS3	F	TGTCACCTACTATCGGGG
	5199–5218	NS3	R	TCAAGCCTATACAGCAAAGG
	5134–5152	NS3	F	CTCCCAGCAAAGATGAACG
	5997–6014	NS4B	R	AGCACCCACACCAACAGC
	5919–5934	NS4B	F	AAGATCTTGAGTGGTG
	6651–6632	NS5A	R	GCCGATAACTCTGACAGC
	6547–6564	NS5A	F	ACACCTGGAAAACAGCCG
	7293–7310	NS5A	R	AGATTCCGTGGCCGAAGG
	7235–7252	NS5A	F	AGCTCTCGTTTCCGGGTG
	7573–7590	NS5B	R	TAGCTGACGCTGTTGTGG
	7511–7527	NS5B	F	ACGCCACCCTATAGGCC
	8027–8046	NS5B	R	GTTGACGGGGAGTGTATTGG
	7926–7943	NS5B	F	ATCGTTTACCCCGATTTG
	8528–8545	NS5B	R	CAAGATGTTATCTGCTCC
	8457–8474	NS5B	F	CGTGACTTCACTAATGCC
	9069–9086	NS5B	R	GTCAATCGAGTTTACGCC
Primer for 5′ RACE	235–252	5′ UTR	F	AATCGCGGCTTGAACGTC
	213–230	5′ UTR	R	TGTACTCACGGATTCACG
Primer for 3′ RACE	8979–8999	NS5B	F	CTTAAAGTACGTGGTGGTCGC
Adapter primer			R	GCGAGCACAGAATTAATACGAC

[a] F, forward; R, reverse.

Ile[176] and Phe[177] were replaced with Ala and Leu (HCVc-mt), respectively, were described previously (28). The gene encoding human signal peptide peptidase (SPP) or its mutant was introduced into pcDNA3.1-myc/His C (Invitrogen) instead of the plasmids described previously (28). The resulting plasmids encoded C-terminally myc-His$_{\times 6}$-tagged wild-type SPP (SPP-wt) or the mutant protein in which Asp[219] was replaced with Ala (SPP-D219A).

Cell culture and transfection. The human embryonic kidney cell line 293FT and the human hepatoma cell line Huh7OK1 (29) were maintained in Dulbecco's modified Eagle's minimal essential medium (DMEM) supplemented with 100 U/ml penicillin, 100 μg/ml streptomycin, nonessential amino acids (Sigma, St. Louis, MO), sodium pyruvate (Sigma), and 10% fetal bovine serum (FBS) and were then cultured at 37°C under the conditions of a humidified atmosphere and 5% CO_2. Plasmids were transfected into cell lines using XtremeGene 8 (Roche) according to the manufacturer's protocol.

Western blot analysis. 293FT cells were cultured in 6-well plates and transfected with the appropriate plasmids. The transfected cells were harvested at 18 h posttransfection, washed with cold phosphate-buffered saline (PBS), and suspended in 50 μl of the lysis buffer consisting of 20 mM Tris-HCl (pH 7.5), 135 mM NaCl, 10% glycerol, 1% Triton X-100, and protease inhibitor cocktail (Merck Bioscience, Calbiochem, San Diego, CA). The lysates were centrifuged at 19,000 × *g* for 5 min at 4°C. The supernatants were mixed with 16 μl of 4× SDS sample buffer and then boiled at 60°C for 20 min. The resulting mixtures were subjected to SDS-PAGE. The proteins in a gel were transferred to polyvinylidene difluoride (PVDF) membranes and incubated with mouse anti-FLAG antibodies (Sigma), mouse anti-HA antibodies (Covance, Princeton, NJ), mouse

anti-c-*myc* antibodies (BD Pharmingen, San Diego, CA), or mouse anti-beta-actin antibodies (Santa Cruz Biotechnology, Santa Cruz, CA) and were then incubated with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies. Immunocomplexes were visualized with SuperSignal West Femto substrate (Thermo Scientific, Rockford, IL) and detected using an LAS-4000 Mini image analyzer (GE Healthcare, Buckinghamshire, United Kingdom).

Detection of antibodies against EHcV. To detect anti-EHcV antibodies in horse sera, we subjected lysates prepared from 293FT cells expressing EHcVc, which is an N-terminally FLAG-tagged and C-terminally HA-tagged EHcV core protein (a positive reference), or cells transfected with an empty plasmid (a negative reference) to Western blotting, as described above. The resulting PVDF membranes were incubated with Blocking One solution (Nacalai Tesque) for blocking at room temperature for 30 min and then incubated with 1,000-fold-diluted horse serum in 10-fold-diluted Blocking One. Mouse anti-FLAG or rabbit anti-EHcV core antibody was used as a positive serum control. The resulting membrane was incubated with an HRP-conjugated antibody to mouse, rabbit, or horse IgG (Abcam, Cambridge, UK) at room temperature for 1 h. Protein bands with a molecular mass of 28 kDa were detected in the positive reference, but not in the negative reference, using positive serum or an antibody to the FLAG epitope tag or EHcV core protein. The rabbit polyclonal antibody against the EHcV core protein was generated by immunization using peptides of the residues from 2 to 15, GNKSKNQKPQPQRG (Scrum Inc., Tokyo, Japan).

Pulldown assay for SPP binding. Human embryonic kidney 293FT cells expressing EHcVc or HCVc with or without SPP-D219A were harvested at 18 h posttransfection, washed with cold PBS, suspended in 100 μl of the lysis buffer, and centrifuged at 14,000 ×g for 5 min at 4°C. Twenty microliters of the lysate was mixed with 20 μl of 2× SDS sample buffer. The remaining lysate was adjusted to 250 μl with the lysis buffer and incubated for 2 h at 4°C after the addition of 20 μl of His-Select nickel affinity gel (Sigma) equilibrated 50% (vol/vol) with lysis buffer. The nickel beads that included SPP-wt or SPP-D219K were washed five times with 500 μl of lysis buffer by centrifugation at 5, 000 × *g* for 1 min at 4°C and then suspended in 40 μl of 1× SDS sample buffer. After being boiled at 60°C for 20 min, the supernatant was subjected to Western blotting to detect the coprecipitated core proteins.

Immunofluorescence microscopy. Huh7OK1 cells were incubated with fresh DMEM containing Bodipy 558/568 (2 μg/ml; Molecular Probes) for 1 h at 37°C to visualize lipid droplets (LDs). The cells were washed once with prewarmed DMEM and incubated for 30 min at 37°C. The treated cells were then fixed in 4% paraformaldehyde for 30 min at room temperature. After two washes with PBS, the cells were permeabilized with permeabilization buffer containing 0.1% saponin (eBioscience, San Diego, CA) for 30 min at 37°C and blocked with PBS containing 2% FBS (blocking buffer) for 30 min at room temperature. The cells were incubated with an appropriate antibody, as indicated in the figure legends. The cells were washed three times with PBS. The mounted cells were observed with a FluoView FV1000 laser scanning confocal microscope (Olympus, Tokyo, Japan). Nuclei were stained with 4′,6′-diamidino-2-phenylindole (DAPI).

Flotation assay. A flotation assay was carried out according to the method described previously (17). Briefly, 293FT cells expressing EHcVc or EHcVc-mt were cultured on a 10-cm dish. The transfected cells were washed once with cold PBS at 18 h posttransfection and harvested using a cell scraper. The cells were suspended in 1.2 ml of 25 mM Tris-HCl (pH 7.5) containing 150 mM NaCl, 5 mM EDTA, and protease inhibitor cocktail (Merck, Calbiochem) (TNE buffer) and were then homogenized by 10 passes through a 26-gauge needle. Each 0.6-ml aliquot of the homogenates was incubated for 30 min on ice with or without 1% Triton X-100 and was then mixed with 0.4 ml of OptiPrep (Axis-Shield, Oslo, Norway). An appropriate concentration of OptiPrep was adjusted with TNE buffer. This mixture was overlaid with 1.2 ml of 30% OptiPrep, 1.2 ml of 25% OptiPrep, and 0.8 ml of 5% OptiPrep, in that order, and was centrifuged at 42,000 rpm for 5 h at 4°C in an SW50.1 rotor (Beckman Coulter, Fullerton, CA). Each fraction, with a volume of 0.4 ml, was collected from the top of the centrifugation tube and was then precipitated by mixing with 4 volumes of cold acetone at −30°C. The resulting pellet was resolved in 50 μl of 1× sample buffer and then subjected to Western blot analysis using a mouse anti-FLAG antibody (Sigma), a rabbit anti-calreticulin antibody (Sigma), and a rabbit anti-caveolin-1 antibody (Sigma). The fractions containing calreticulin in the absence and presence of Triton X-100 were defined as the membrane and detergent-soluble membrane fractions, respectively. In the presence of the detergent, fractions 3 to 5, which contained caveolin-1 but only small amounts of calreticulin, were defined as the detergent-resistant membrane fractions.

Nucleotide sequence accession numbers. The whole sequence of the EHcV strain isolated from serum sample 3 was deposited in GenBank under accession number AB863589. The nucleotide sequences of the partial NS5B-coding regions and 3′ UTRs from samples 5 and 1 were registered as AB921150 and AB921151, respectively.

RESULTS

Detection of the EHcV genome and antibody to EHcV in sera of Japanese-born horses. To clarify whether NPHVs were distributed in Japan, we collected 31 horse serum samples and examined them in order to detect the EHcV genome and antibody to the core protein. We prepared total RNAs from horse sera and screened them using RT-PCR analyses with three sets of PCR primers (NPHV-F1/NPHV-R1, NPHV-F2/NPHV-R2, and NPHV-F3/NPHV-R3) that targeted the NS3-coding region that is relatively conserved among NPHVs. Total RNA prepared from conventional rabbit serum was used as a negative control. PCR products with the expected sizes were found in horse serum samples 1, 3, 25, 26, 27, and 29 to 31 using NPHV-F1/NPHV-R1, in horse serum samples 3 and 5 using NPHV-F2/R2, and in horse serum samples 1, 3, 5, 20, and 25 to 31 using NPHV-F3/R3 (Fig. 1A). The EHcV genome was detected in 11 of 31 (35%) serum samples by RT-PCR (Fig. 1A and B). Copy numbers of the EHcV genome in horse sera varied from 10^4 to 10^9 copies per ml of sera (Fig. 1B). Although a PCR product was slightly amplified from serum sample 19 by PCR using the primer pair NPHV-F1/R1, the copy number of the virus genome in serum sample 19 was estimated to be low, at a level similar to that of the negative control. Thus, we could not determine whether serum sample 19 included a viral genome. We then immunologically surveyed horse sera by Western blotting. Western blotting analyses using horse sera to detect antibodies to the EHcV core protein (Fig. 2) showed that the sera of samples 1, 2, 3, 5, 14, 20, and 25 were immunoreactive to the EHcV core protein (7 positive serum samples of a total of 31 samples; 22.6%). The sera of samples 1, 3, 5, and 20 were PCR positive and seropositive. Serum samples 2 and 14 were PCR negative and seropositive, whereas samples 26 to 31 were PCR positive and seronegative. These results suggest that EHcV has infected Japanese-born domestic horses.

Genetic analysis of EHcV. PCR products corresponding to the 5′ UTR and the open reading frame were segmentally amplified from serum sample 3 by 5′ RACE and RT-PCR, respectively. In the present study, we successfully determined the 3′-terminal sequence downstream of a stop codon using the 3′-RACE method with poly(U) polymerase. We determined the nucleotide sequence of the putative full genome, which was designated JPN3/JAPAN/2013 (GenBank accession number AB863589). The full-length genome of strain JPN3/JAPAN/2013 is composed of 9,355 nucleotides, consisting of the 5′ UTR with a nucleotide length of 389, the 3′ UTR with a nucleotide length of 134, and an open

Downloaded from http://jvi.asm.org/ on February 15, 2015 by guest

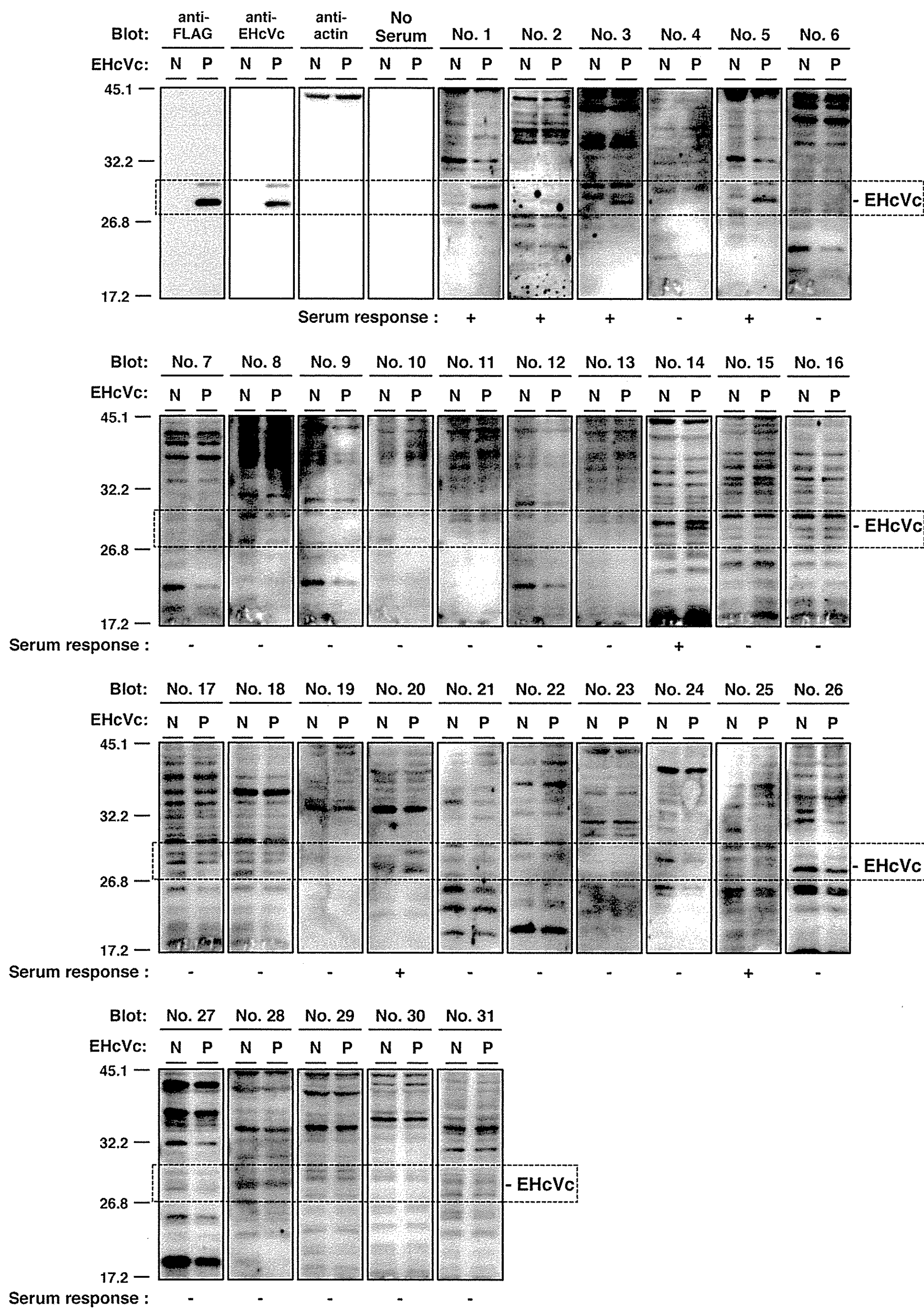

FIG 2 Serological screening of Japanese-born domestic horses. Lysates of 293FT cells transfected with an empty plasmid (a negative reference, N) or the plasmid encoding EHcVc (a positive reference, P) were subjected to Western blotting using serum from each horse. The serum response "+" indicates that the protein band with the same molecular size as the EHcV core protein was specifically detected in the "P" lane, but not in the "N" lane, while the serum response "−" indicates that the protein band with the same molecular size as the EHcV core protein was detected in neither the "P" lane nor the "N" lane. Both antibodies to the FLAG tag and to the EHcV core protein were used as serum positive controls, while protein amounts were standardized with blotting using the antibody to beta-actin. "No serum" indicates the membrane was incubated without primary antibodies but with HRP-conjugated anti-horse IgG antibodies as a background of the secondary antibody.

reading frame with a nucleotide length of 8,832. The open reading frame encodes 2,943 amino acids. Table 2 summarizes the amino acid homology of the JPN3/JAPAN/2013 polyprotein with the polyproteins of the other EHcV strains. The polyprotein of JPN3/JAPAN/2013 shared more than 94% homology with the other EHcV polyproteins and exhibited the highest homology, 97.8%, with NPHV-H10-094 (GenBank accession number JQ434007), which was isolated from a horse in the United States (8). The NS3- and NS5B-coding regions of the EHcV strains were phylogenetically analyzed by the neighbor-joining method. The phylogenetic

TABLE 2 Amino acid sequence homologies of the polyproteins

	Non-primate hepaciviruses					
	H10-094 (JQ434007)	B10-022 (JQ434004)	NZP1 (JQ434001)	AAK-2011 (JF44991)	H3-011 (JQ434008)	A6-066 (JQ434003)
JPN3/JAPAN/2013 (AB863589)	97.8[a]	96.7	95.7	95.7	95.6	95.3

	Non-primate hepaciviruses		HCV			
	G1-073 (JQ434002)	F8-068 (JQ434005)	HCV1a (NC004102)	HCV1b (AB779562)	JFH1 (AB047639)	GBV-B (NC001655)
JPN3/JAPAN/2013 (AB863589)	94.9	94.1	46.5	45.6	44.5	28.9

	Non-primate hepaciviruses		
	JPN3/JAPAN/2013 (AB863589)	AAK-2011 (JF44991)	GBV-B (NC001655)
HCV1a (NC004102)	46.1	46.0	33.3

[a], percent identity.

trees of the NS3 (Fig. 3A) and NS5B regions (Fig. 3B) showed that JPN3/JAPAN/2013 was included in the clade comprising the U.S. strains NPHV-H10-094 (GenBank accession number JQ434007) and B10-022 (GenBank accession number JQ434004).

Putative RNA secondary structures around the UTRs of EHcV. The 5′-terminal region of JPN3/JAPAN/2013 was compared with those of the EHcV genomes (Fig. 4A). The HCV internal ribosome entry site (IRES)-like structure was embedded in the 5′ UTRs of NPHVs (5, 6). The 5′-UTR region was well conserved among the EHcV strains and showed a mean diversity of approximately 4% among the EHcV strains (Fig. 4A). The 3′-terminal sequence downstream of the (A)-rich region in the EHcV genome had not been reported because the (A)-rich region downstream of the stop codon of EHcV interrupted the reaction in the ordinary 3′-RACE method (5, 6). In the present study, we determined the nucleotide sequences downstream of the (A)-rich region from serum sample 3 (JPN3/JAPAN/2013; GenBank accession number AB863589), sample 5 (JPN5/JAPAN/2015; GenBank accession number AB921150), and sample 1 (JPN1/JAPAN/2015; GenBank accession number AB921151) by the modified 3′-RACE method using poly(U) polymerase, although the region in serum sample 1 was incompletely amplified (Fig. 4B). The regions downstream of the (A)-rich region were conserved between serum samples 3 and 5, whereas the (A)-rich regions varied among the three strains (Fig. 4B).

The secondary structure of 5′ UTR in strain JPN3/JAPAN/2013 was predicted according to the method described previously (8) (Fig. 4C). The stem-loops in the 5′ UTR were designated according to the stem-loops of the HCV 5′-UTR structures (30). Stem-loops (SLs) I, II, IIIa to IIIf, and the pseudoknot interaction were predicted within the 5′ UTR of strain JPN3/JAPAN/2013. These structures were the same as that of the strain reported previously (9), although several nucleotide insertions and deletions were more predominant in the apical loop of subdomain IIIb than in the other strains reported previously (Fig. 4A and C). Two seed sites of the microRNA miR-122 (Fig. 4A and C) were found in the 5′ UTR of strain JPN3/JAPAN/2013 at nucleotide residues 81 to 89 (UCCACAUUA) and 98 to 103 (CACUCC), which also corresponded to the predicted miR-122 seed sites in the 5′ UTRs of the other EHcV strains (9).

The HCV 3′ UTR, which is generally 200 to 300 nucleotides in length, consists of a short variable region, the poly(U/UC) stretch sequence, and the 3′-X-tail region, in that order (31–33). Although the EHcV 3′ UTR, which is composed of 138 nucleotides, is shorter than the HCV 3′ UTR, the 3′ UTR of EHcV consists of the (A)-rich sequence and 3′-X-tail region, in that order. The (A)-rich sequence of EHcV may vary in length (Fig. 4B). We subsequently predicted the secondary structure of the EHcV 3′ UTR. Although the EHcV 3′ UTR, which is composed of 138 nucleotides, is shorter than the HCV 3′ UTR, the 3′ UTR includes three predicted SL structures (Fig. 4C). Based on the SL structures in the HCV 3′ X-tail, these SL structures in the EHcV 3′ UTR were designated 3′SL I, 3′SL II, and 3′SL III, in that order from the 3′ terminus (Fig. 4C). Interestingly, the (A)-rich sequence was partially incorporated into the 3′SL III, although the poly(U/UC) stretch sequence in the HCV 3′ UTR is separated from any 3′SL structures (31–33). Furthermore, the two SL structures in the 3′ side of the EHcV NS5B-coding region were predicted to correspond to 5BSL3.2 and 5BSL3.3 in the NS5B-coding region of HCV. HCV 5BSL3.2 was previously shown to interact with 3′SL II to form the kissing-loop interaction, which is required for HCV replication (33). The secondary structure prediction shown in Fig. 4C suggests that the kissing-loop interaction may be conserved between 5BSL3.2 and the 3′SL II of the EHcV genome through their complementary sequences. The long-range RNA-RNA interaction between the apical loop of subdomain IIId in HCV IRES and the bulge of 5BSL3.2 supports IRES-dependent translation and viral RNA replication (34–36). In the case of the EHcV genome, the complement sequences were detected in the apical loops of subdomain and the 5BSL3.2-like subdomain (Fig. 4C), suggesting that the long-range RNA-RNA interaction may reside in the EHcV genome. These results indicated that HCV-like RNA secondary structures may be conserved around both UTRs of the EHcV genome.

Cleavage of the EHcV core protein by SPP. The C-terminal transmembrane region of the HCV core protein was previously shown to be cleaved by SPP following the cleavage of the core-E1 junction by signal peptidase (11, 28, 37). The core protein is known to be released from the precursor polyprotein embedded in the endoplasmic reticulum (ER) membrane, and it then moves

Downloaded from http://jvi.asm.org/ on February 15, 2015 by guest

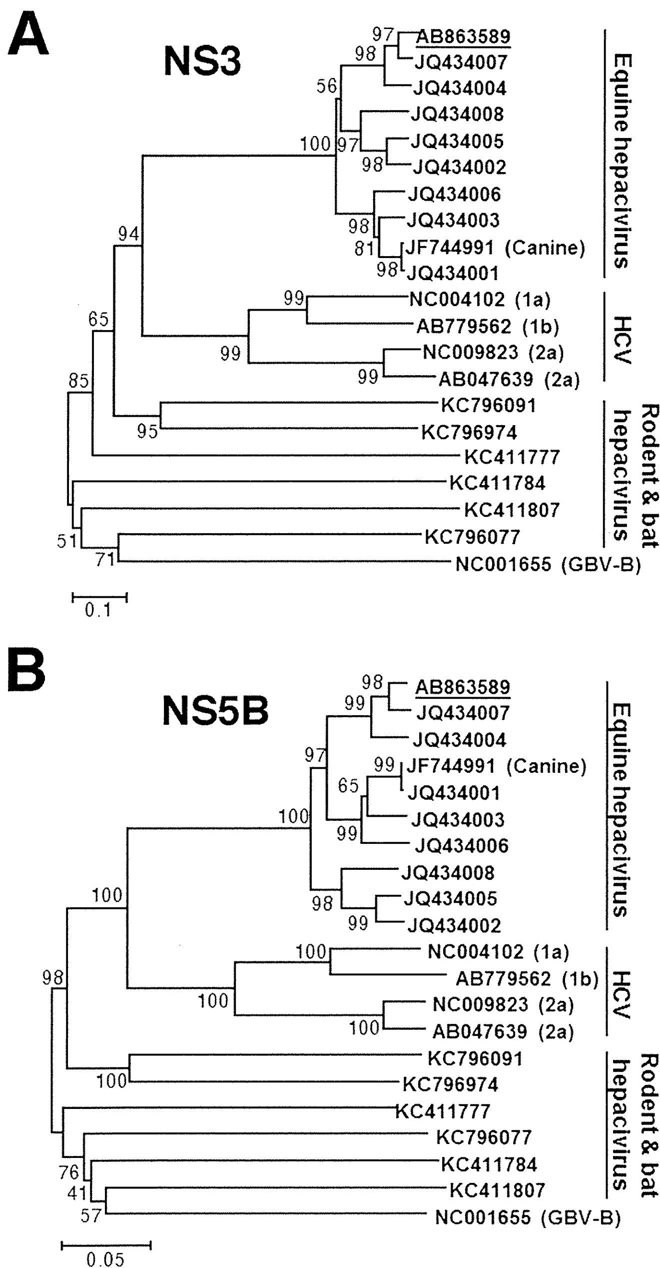

FIG 3 Phylogenetic analysis of the EHcV gene. Neighbor-joining trees of the nucleotide sequences from the NS3 (A) and NS5B (B) regions of the NPHV, HCV, and GBV-B strains are shown (23). Trees were constructed by the maximum composite likelihood method calculated using the program MEGA5 (24). The percentage of replicate trees in which the associated taxa were clustered together in the bootstrap test (1,000 replicates) is indicated next to the branches. Analyses were carried out using 10 strains of EhcV, JPN3/JAPAN/2013, A6-066 (GenBank accession no. JQ434003), B10-022 (GenBank accession no. JQ434004), F8-068 (GenBank accession no. JQ434005), G1-073 (GenBank accession no. JQ434002), G5-077 (GenBank accession no. JQ434006), H3-011 (GenBank accession no. JQ434008), H10-094 (GenBank accession no. JQ434007), NZP1 (GenBank accession no. JQ434001), and AAK-2011 (canine hepacivirus; GenBank accession no. JF744991); 4 strains of HCV, H77 (genotype 1a; GenBank accession no. NC004102), LyHCV (genotype 1b; GenBank accession no. AB779562), HC-J6CH (genotype 2a; GenBank accession no. NC009823), and JFH1 (genotype 2a; GenBank accession no. AB047639); 3 strains of bat hepacivirus, PDB-112 (GenBank accession no. KC796077), PDB-445 (GenBank accession no. KC796091), and PDB-829 (GenBank accession no. KC796074); 3 strains of rodent hepacivirus, RMU10-3382 (GenBank accession no. KC411777), NLR-AP-70 (GenBank accession no. KC411784), and SAR-46 (GenBank accession no. KC411807); and another primate hepacivirus, GBV-B (GenBank accession no. NC001655). The Japanese strain JPN3/JAPAN/2013 (GenBank accession no. AB863589) is underlined.

mainly to lipid droplets (LDs) (13, 14). Although SPP-dependent cleavage and LD translocation of the capsid protein are features common to HCV and GBV-B (13), it currently remains unknown whether the EHcV core protein shows these properties. The EHcV core protein shared 49.5% amino acid homology with the HCV core protein (genotype 1b) (Fig. 5A) and exhibited a hydrophobic/hydrophilic pattern similar to that of the HCV core protein (Fig. 5B). The EHcV core protein was predicted to be composed of domains 1, 2, and 3 relative to the HCV core protein. The transmembrane region of the EHcV core protein was predicted to span from Asn^{177} to Val^{199} by TMHMM2.0 (http://www.cbs.dtu.dk/services/TMHMM/). The transmembrane region of the EHcV core protein was 65% identical to that of the HCV core proteins (Fig. 5A). The C-terminal residue of the mature HCV core protein was found to be Phe^{177} in human and insect cell lines (17, 38). Our previous findings suggest that Ile^{176} and Phe^{177} of the HCV core protein may be responsible for SPP-dependent cleavage, because the replacement of Ile^{176} and Phe^{177} with Ala and Leu, respectively, abrogated intramembrane cleavage by SPP and impaired virus production (17, 28, 39). Weihofen et al. reported that SPP cleaved a peptide bond of the alpha-helix-breaking structure in a transmembrane region of the membrane protein (40). The replacement of Ile^{176} and Phe^{177} with Ala and Leu, respectively, in the HCV core protein converted the beta-sheet structure (alpha-helix-breaking structure) to an alpha-helix structure in the transmembrane region, as reported previously (28) (Fig. 6A and B). Ile^{190} and Phe^{191} of the EHcV core protein, which correspond to Ile^{176} and Phe^{177}, respectively, of the HCV core protein, reside in the alpha-helix-breaking structure of the transmembrane region (Fig. 6A and B). In contrast, the replacement of Ile^{190} and Phe^{191} with Ala and Leu, respectively, in the EHcV core protein were predicted to convert the beta-sheet to an alpha-helix structure in a manner similar to that for the HCV core protein (Fig. 6A and B). To investigate the involvement of SPP in the maturation of the EHcV core protein, we expressed EHcVc or HCVc in 293FT cells with an SPP or SPP mutant. These core proteins were expected to be resistant to signal peptidase-dependent processing because the C-terminal residue Ala of both core proteins was replaced with Arg, resulting in the detection of an immature core protein by the anti-HA antibody (Fig. 6A) (28). The core proteins with molecular masses of 23 kDa and 28 kDa were detected mainly with the anti-FLAG antibody in 293FT cells expressing HCVc and HCVc-mt, respectively (Fig. 6C, lanes 2 and 3); however, the 23-kDa band was not detected with the anti-HA antibody (Fig. 6C, lane 2). When EHcVc was expressed in 293FT cells, it was detected at a molecular mass of 27 kDa with the anti-FLAG antibody, but not with the anti-HA antibody (Fig. 6C, lane 4). In contrast, EHcV-mt, in which the 190th and 191st residues were Ala and Leu instead of Ile and Phe, respectively, was detected mainly at a molecular mass of 30 kDa with the anti-FLAG and anti-HA antibodies (Fig. 6C, lane 5). A loss-of-function SPP mutant (SPP-D219A) in which the 219th residue was Ala instead of Asp was shown to have a dominantly negative effect on SPP-dependent cleavage of the

A

```
           1                                         I
AB863589: ACCUCCGUGC UAUGCACGGU GCGUUGUCAG CGUUUUGCGC UUGCAUGCGC UACACGCGUC GUCCAACGCG GAGGGAUUCU UCCACAUUAC CAUGUGUCAC
JQ434007: ---------- ---GCACGGU GCGUUGUCAG CGUUUUGCGC UUGCAUGCGC UACACGCGUC GUCCAACGCG GAGGGAU--U UUCACAUUAC CAUGUGUCAC
JQ434003: ACCUCCGUGC UUGGCACGGU GCGUUGUCAG CGUUUUGCGC CUGCAUGCGC UACACGCGUC GUCCAACGCG G-GGGAG--U UCCACACUAC CAUGUGUCAC
JQ434008: ACCUCCGUGC UAGGCACGGU GCGUUGUCAG CGUAUUGCGC UUGCAUGCGC UCCACGCGUC GUCCAACGCG GAGGGAA--U UUCACAUUAC CAUGUGUCAC
JQ434005: ACCUCCGUGC UUGGCACGGU GCGUUGUCAG CGUGUUGCGC UUGCAUGCGC UCUACGCGUC GUCCAACGCG GAGGGAA--U UUCACAUUAC CAUGUGUCAC
JQ434002: ACCUCCGUGC UAGGCACGGU GCGUUGUCAG CGUAUUGCGC UUGCAUGCGC UCCACGCGUC GUCCAACGCG GAGGGAA--C UUCACAUUAC CAUGUGUCAC
          .......... .*........ .......... ...*...... *......... .*........ .......... ......*... .......... ..........
          101                                         II
AB863589: UCCCCCUAUG GAGGGUUCCA CCCCGCCCAC ACGGAAAUAG GUUAACCAUA CCUAUAGUAC GGGUGAGCGG GUCCUCCUAG GGCCCCCCCG GCAGGUCGAG
JQ434007: UCCCCCUAUG GAGGGUUCCA CCCCGCCUAC ACGGAAAUGG GUUAACCAUA CCCAUAGUAC GGGUGAGCGG GUCCUCCUAG GGCCCCCCCG GCAGGUCGAG
JQ434003: UCCCCCUAUG GAGGGUUCCA CCCCGCUUAC ACGGAAAUGG GUUAACCACA CCCAUAGUAC GGGUAUGCGG GUCCUCCUAG GGCCCCCCCG GCAGGUCGAG
JQ434008: UCCCCCUGUG GAGGAACCCA -CCCGUUUAC ACGGAAAUGG GUUAACCAUA CCCAAAGUAC GGGUACUCGG GUCCUCCCAG GGCCCCCCCG GCAGGUCGAG
JQ434005: UCCCCCUAUG GAGGGAUCCA -CCCGUCUAC ACGGAAACGG GUUAACCAUA CUCUGAGUAC GGGUAUUCGG GUCCUCCUAG GGCCCCCCCG GCAGGUCGAG
JQ434002: UCCCCCUAUG GAGGGAUCCA -CCCGUCUAC ACGGAAAUGG GUUAACCAUA CCCAGAGUAC GGGUAUUCGG GUCCUCCUAG GGCCCCCCCG GCAGGUCGAG
          .......... .....*.... .....*.... .......... .......... ...**..... ....***... .......... .......... ..........
          201                     IIIa                              IIIb                        IIIc
AB863589: GGAGCUGAAA UUCGUGAAUC CGUGAGUACA CGGAAAUCGC GGCUUGAACG UCAUA-CGUG ACCUUCGGAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
JQ434007: GGAGCUGAAA UUCGUGAAUC CGUGAGUACA CGGAAAUCGC GGCUUGAACG UCAUA-CGUG ACCUUCGGAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
JQ434003: GGAGCUGGAA UUCGUGAAUU CGUGAGUACA CGAAAAUCGC GGCUUGAACG UCUU----UG ACCUUCGGAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
JQ434008: GGAGCUGGAA UUCGUGAAUC CGUGAGUACA CGGAAAUCGC GGCUUGAACG UCUUUUAAUG ACCUUCGGAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
JQ434005: GGAGCUGGAA UUCGUGAAUC CGUGAGUACA CGGAAAUCGC GGCUUGAACG UCUUA---GG ACCUUCGGAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
JQ434002: GGAGCUGGAA UUCGUGAAUC UGUGAGUACA CGGAAAUCGC GGCUUGAACG UCUUU---GG ACCUUCGAAG CCGAAAUUUG GGCGUGCCCC ACGAAGGAAG
          .......... .......... .......... .......... .......... ....*.***. .......... .......... .......... ..........
          301         IIId                                   IIIe        IIIf
AB863589: GCGGGGGCGG UGUUGGGCCG CCGCCCCCUU UAUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
JQ434007: GCGGGGGCGG UGUUGGGCCG CCGCCCCCUU UAUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
JQ434003: GCGGGGGCGG UGUUGGGCCG CCGCCCCCUU UAUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
JQ434008: GCGGGGGCGG UGUUGGGCCG CCGCCCCCUU UAUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
JQ434005: GCGGGGGCGG UGUUGGGCCG CCGCCCCCCU AGUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
JQ434002: GCGGGGGCGG UGUUGGGCCG CCGCCCCCCU UGUCCCACGG UCUGAUAGGA UGCUUGCGAG GGCACCUGCC GGUCUCGUAG ACCAUAGGAC
          .......... .......... .......... **........ .......... .......... .......... ....... .......... ..
```

B

```
          9141
AB863589: CUGCCGAUCG UUGCUGGGAG GUGGAUUCCU CUUCUUCUCU UGCUUCCUUU ACUGGCUGCA AUUCUUUACU UCAAUAAAUA A--------- -------AAA
AB921150: CUUCCGGUCG UUGCCGGGAA GUGGAUACCU CUUCUGCUUU UGUUGCCUCU CUUGGCUGCA AUUCUUUAUU UCAAUAAAUA GAUAAAUAAA UCAAAAAAAA
AB921151: CUUCCGGUCG UUGCCGGGAG GUGGAUACCU CUCUUGUUUU UGUUACCUCU UUUGGCUGCA AUUCUUUACU UUAACAAAUA A------AAC CUAAAAAAAA
                                             NS5B coding region
          9225
AB863589: AAAAAAG--- ---AAAAUAA UUA---AAAU UCUUUAGACU UCCUUUCCCC UUCCCCUUUU UCCCGUUUAG GUUACUUCCU AUGGAAGAAC AGGAGGGUGG
AB921150: AAAAAAA--- ---AAAAUAA UUA---AA-- ----UAGUUU UCCUUUCCCU UUUCCCCUUU UCCCUUU--G GUUACUUCCU AUGGAAGAAC AGGAGGGUGG
AB921151: AAAAAAACCA AUCAAAA--- ----UAAUAA GAUAAAUUAA UCAGUUUCCC CUUUCCUCCC CCUC------ ---------- ---------- ----------
                  (A)-rich region                                                  3' X-tail
          9316                           9355
AB863589: GUGUAUGGGA GCCCUGUUUG GCCCCUAUGG GGCCAAGUUG
AB921150: GUG-ACGGGA GCCCUGUCCU GCCCCUAUGG GGCAGCUUUA
AB921151: ---------- ---------- ---------- ----------
```

C

FIG 4 RNA structure analysis of the 5′ and 3′ ends of EHcV. (A) Alignment of the 5′ UTRs of the EHcV strains. Asterisks and black lines indicate variable residues and miRNA-targeting regions, respectively. SL structures are enclosed in rectangles. GenBank numbers are listed in the legend to Fig. 3. (B) Alignment of the 3′ UTRs of the EHcV strains. The 3′-terminal sequences with GenBank numbers AB921150 and AB921151 were determined using serum samples 5 and 1, respectively, and were aligned with that of JPN3/JAPAN/2013 (GenBank accession no. AB863589). (C) The secondary structures of the 5′ UTR, NS5B-coding region, and 3′ UTR were predicted on the basis of minimum free energy predictions. The stem loops in the 5′ UTR (left) were designated according to the stem loops of the HCV 5′-UTR structures (30). A gray line, double line, and single line indicate a pseudoknot structure, miR-122-target region, and complementary sequence, respectively. The stem loops in the NS5B-coding region and 3′ UTR (right) were designated to correspond to the stem loops embedded in the HCV 3′-terminal region (32, 33). The NS5B-coding region, (A)-rich region, and 3′-X-tail sequence are indicated under the schematic structure. Start and stop codons are enclosed by rectangles.

Downloaded from http://jvi.asm.org/ on February 15, 2015 by guest

A

```
                                     Domain 1
EHcV      1: MGNKSKNQKP QPQRGPRNRV KGQSRSGPVV FPSGAVLVGG RYIPPPKKAI RGPRRGLVQA PKSSERTAPR
HCV 1b    1: M---STN--P KPQRKTKRN- -TNRRPQDVK FPGGGQIVGG VY-LLP---R RGPRLGVRAT RKTSERSQPR

                                     Domain 1
EHcV     71: KRRQPPPQTD SSWRKYFPKF WGDKGYPWP- YVDPVLQWGA WGSSPGAYRT RWGPRDPRHR SRNLGRVIDT
HCV 1b   60: GRRQPIP--- -KARRPEGRA WAQPGYPWPL YGNEGLGWAG WLLSPRGSRP SWGPTDPRRR SRNLGKVIDT

                          Domain 2                                  Domain 3
EHcV    140: LTCGVADLAG YVPVLGAPAG ALCRGAAHLV RFVEDGANFI TGNIPGMGFS IFLLALFS-A VSFGEASVVR
HCV 1b  126: LTCGFADLMG YIPLVGAPLG GAARALAHGV RVLEDGVNYA TGNLPGCSFS IFLLALLSCL TIPASAYEVR

                                                          Transmembrane region
```

B

Domain 1 | Domain 2 | Domain 3 — EHcV core protein

Domain 1 | Domain 2 | Domain 3 — HCV 1b core protein

FIG 5 Amino acid alignment and hydrophobicity of EHcV and HCV core proteins. (A) Alignment of the core proteins of EHcV (JPN3/JAPAN/2013) and HCV genotype1b (Con1; GenBank accession number AJ238799). Asterisks indicate identical amino acid residues. Bars indicate gaps to achieve maximum amino acid matching. The black and white arrowheads indicate the predicted cleavage site of the core protein of HCV by SPP and signal peptidase, respectively. The EHcV core protein was composed of three domains, domain 1 (a black line, residues 2 to 132), domain 2 (a broken black line, residues 133 to 187), and domain 3 (a gray line, residues 188 to 204), relative to those of the HCV core protein (42). (B) Hydrophobicity plots of the EHcV and HCV core proteins were prepared by the method of Kyte and Doolittle (26). The horizontal and vertical axes represent amino acid position and hydrophobicity, respectively.

HCV core protein and to be coprecipitated with the immature core protein (28). HCVc had a molecular mass of 23 kDa in the presence of wild-type SPP (SPP-wt) and was detected with the anti-FLAG antibody, but not with the anti-HA antibody (Fig. 6D, lane 2), suggesting that the 23-kDa protein band may be a mature core protein. HCVc mainly had a molecular mass of 28 kDa in the presence of SPP-D219A and was detected with the anti-FLAG and anti-HA antibodies (Fig. 6D, lane 3), which corresponds to the mobility of HCVc-mt (Fig. 6D, lane 4). These results suggest that SPP-D219A may abrogate the intramembrane cleavage of HCVc. In a manner similar to that for the HCV core protein, EHcVc was detected mainly at a molecular mass of 27 kDa in the presence of SPP-wt with the anti-FLAG antibody, but not with the anti-HA antibody (Fig. 6D, lane 6). EHcVc was detected mainly at a molecular mass of 30 kDa in the presence of SPP-D219A with anti-FLAG and anti-HA antibodies (Fig. 6D, lane 7), corresponding to the mobility of EHcVc-mt (Fig. 6D, lane 8). When SPP-D219A was coexpressed with either HCVc or EHcVc, immature HCVc and EHcVc were coprecipitated with SPP-D219A (Fig. 6E, lanes 3 and 5). These results suggest that the EHcV core protein may be cleaved by SPP and that Ile^{190} and Phe^{191} of the EHcV core protein are critical for SPP-dependent cleavage.

The intracellular localization of the hepacivirus core protein. The HCV core protein is known to be localized mainly on the surface of LDs and is partially fractionated in the detergent-resistant membrane (DRM) close to the budding sites on the ER (13, 14, 16, 17). The core protein is considered to encompass the viral genome on the ER membrane, followed by budding into the lumen side (13, 14, 16, 17). To examine the intracellular localization of the EHcV core protein, we expressed HCVc or EHcVc in the Huh7OK1 cell line and stained the core proteins with the anti-FLAG antibody after staining LDs. Consistent with the findings of previous studies (14, 41, 42), HCVc was localized mainly on LDs (Fig. 7, row 3), whereas HCV-mt was not (Fig. 7, row 4). In a manner similar to that for the HCV core protein, EHcVc was localized mainly on LDs (Fig. 7, top row), whereas EHcVc-mt was not (Fig. 7, second row). These results suggest that the EHcV core protein may be localized mainly on LDs after SPP-dependent cleavage.

The DRM is defined as the cholesterol/sphingolipid-rich microdomain, which is resistant to nonionic detergents such as Triton X-100, considered to be a characteristic of lipid rafts. HCV was previously shown to be propagated in lipid raft-like compartments, including the membranous web (43–45). Furthermore, the HCV core protein is known to be associated with lipid raft-like compartments as well as LDs (16, 17, 41, 42). Therefore, we determined whether the EHcV core protein could be detected in the DRM fractions. EHcVc or EHcVc-mt was expressed in 293FT cells. The resulting cells were lysed on ice in the presence or absence of 1% Triton X-100. The DRM fractions were separated from the soluble proteins by a flotation assay with a stepwise density gradient in the presence or absence of Triton X-100. Serial fractions were collected after ultracentrifugation and were then subjected to Western blot analysis after being concentrated. EHcVc and EHcVc-mt were fractionated broadly from fractions of samples 3 to 11 without Triton X-100, and the fraction from

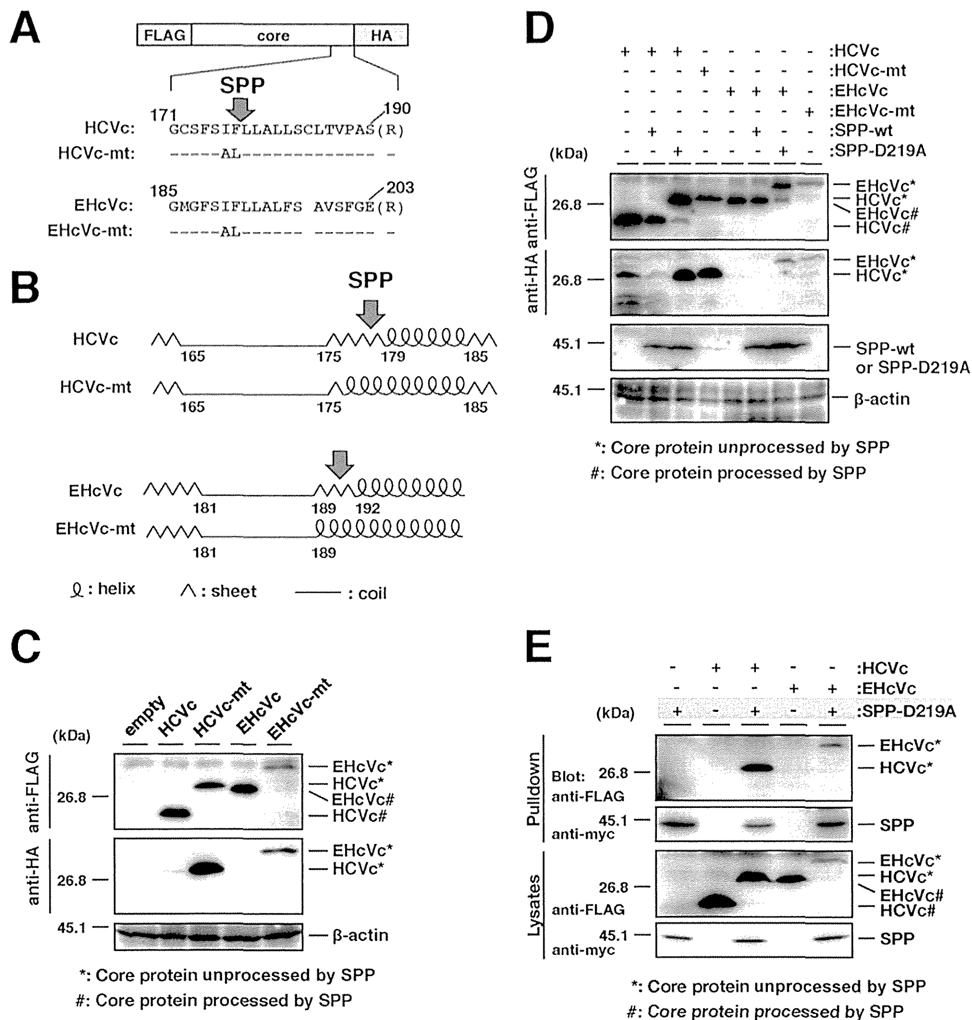

FIG 6 Intramembrane processing of the EHcV core protein by SPP. (A) The plasmids encoding HCVc, HCVc-mt, EHcVc, and EHcVc-mt are shown as a schematic diagram. Their C-terminal regions (171 to 190, HCV core protein; 185 to 203, EHcV core protein) were aligned. The C-terminal Ala of each core protein was replaced with Arg (R) to prevent signal peptidase-dependent cleavage for the detection of the SPP-uncleaved core protein with the anti-HA antibody. Bars indicate the amino acids that were the same as those of the wild-type residues. (B) The secondary protein structures in the C-terminal transmembrane regions of the HCV and EHcV core proteins and mutants were predicted by the method of Garnier et al. (25). Arrows indicate putative SPP cleavage sites. (C) HCVc, HCVc-mt, EHcVc, and EHcVc-mt were expressed in 293FT cells and immunoblotted with the anti-FLAG and -HA antibodies. (D) HCVc or EHcVc was expressed with SPP-wt or SPP-D219A in the 293FT cell line. HCVc-mt and EHcVc-mt were expressed in the absence of SPP-wt and SPP-D219A as uncleavable controls. (E) HCVc or EHcVc was coexpressed with or without SPP-D219A. SPP-D219A was pulled down with Ni beads. Coprecipitated proteins were immunoblotted with the anti-FLAG antibody.

sample 8 contained the largest amount of the core protein (Fig. 8, left panels). The distributions of the core proteins were roughly consistent with that of calreticulin, a marker protein of the ER membrane. When the cells expressing EHcVc were lysed in the presence of Triton X-100, a large amount of the core protein was localized in fractions 9 to 11 (Fig. 8, top three panels on the right). These fractions were enriched in calreticulin, corresponding to the detergent-soluble fractions (Fig. 8, fractions 7 to 11, top three panels on the right). However, EHcVc was partially detected in fractions 3 to 6 together with caveolin-1, a marker protein of the lipid raft (Fig. 8, fractions 3 to 6, top three panels on the right), suggesting that the EHcV core protein may have been partially distributed in the DRM fractions. In contrast, EHcVc-mt was localized in the detergent-soluble fractions (Fig. 8, fractions 9 to 11, bottom three panels on the right), but not in the DRM fractions (Fig. 8, fractions 3 to 6, bottom three panels on the right), in the presence of Triton X-100. EHcVc-mt was resistant to SPP-dependent processing, as described above (Fig. 6). These results suggest that the EHcV core protein may have been partially localized in the DRM and also that SPP-dependent processing may be required for DRM localization of the EHcV core protein.

DISCUSSION

The results of the present study indicate that EHcV infects Japanese-born domestic horses. Previous studies suggested that EHcV infected mainly horses and rarely dogs (5, 7–9). Our results demonstrate that EHcV commonly infects Japanese-born domestic horses (35.6% PCR positive and 22.6% seropositive). Several groups reported a prevalence of less than 10% PCR positivity in horses raised in the United States, the United Kingdom, and Ger-

Downloaded from http://jvi.asm.org/ on February 15, 2015 by guest