nature genetics

A variant upstream of *IFNL3* (*IL28B*) creating a new interferon gene *IFNL4* is associated with impaired clearance of hepatitis C virus

Ludmila Prokunina-Olsson[1], Brian Muchmore[1], Wei Tang[1], Ruth M Pfeiffer[2], Heiyoung Park[3], Harold Dickensheets[4], Dianna Hergott[1,5], Patricia Porter-Gill[1], Adam Mumy[1], Indu Kohaar[1], Sabrina Chen[6], Nathan Brand[1], McAnthony Tarway[1], Luyang Liu[1], Faruk Sheikh[4], Jacquie Astemborski[7], Herbert L Bonkovsky[8], Brian R Edlin[9,10], Charles D Howell[11], Timothy R Morgan[12,13], David L Thomas[7,14], Barbara Rehermann[3], Raymond P Donnelly[4] & Thomas R O'Brien[5]

Chronic infection with hepatitis C virus (HCV) is a common cause of liver cirrhosis and cancer. We performed RNA sequencing in primary human hepatocytes activated with synthetic double-stranded RNA to mimic HCV infection. Upstream of *IFNL3* (*IL28B*) on chromosome 19q13.13, we discovered a new transiently induced region that harbors a dinucleotide variant ss469415590 (TT or ΔG), which is in high linkage disequilibrium with rs12979860, a genetic marker strongly associated with HCV clearance. ss469415590[ΔG] is a frameshift variant that creates a novel gene, designated *IFNL4*, encoding the interferon-λ4 protein (IFNL4), which is moderately similar to IFNL3. Compared to rs12979860, ss469415590 is more strongly associated with HCV clearance in individuals of African ancestry, although it provides comparable information in Europeans and Asians. Transient overexpression of IFNL4 in a hepatoma cell line induced STAT1 and STAT2 phosphorylation and the expression of interferon-stimulated genes. Our findings provide new insights into the genetic regulation of HCV clearance and its clinical management.

More than 3% of the world population is infected with HCV[1]. Up to 80% of acutely infected individuals fail to clear the virus and develop chronic hepatitis C (CHC)[2], with as many as 5% eventually progressing to liver cancer[3]. Success of CHC treatment with pegylated interferon-α with ribavirin (pegIFN-α/RBV) depends on HCV genotype and reaches 50–80% in patients of European ancestry but only 30% in patients of African ancestry. Adding a direct-acting antiviral (DAA) agent to this regimen increases the success rate but may result in adverse effects from pegIFN-α/RBV[4] and DAA treatment. If this treatment fails, there is an increased risk of the selection of resistant HCV strains that may compromise future treatment options[5,6].

Recent genome-wide association studies (GWAS) have identified the rs12979860 and rs8099917 SNPs on chromosome 19q13.13 near the *IFNL3* gene (formerly known as *IL28B*) as variants associated with both spontaneous HCV clearance[7,8] and response to pegIFN-α/RBV treatment[7–11]. Within this region reside the three interferon-λ genes, *IFNL1*, *IFNL2* and *IFNL3* (formerly *IL29*, *IL28A* and *IL28B*, respectively), which encode the type III IFNs[12,13]. Type I IFNs (IFN-α and IFN-β) and type III IFNs (IFN-λs) induce antiviral activity and suppress HCV replication *in vitro*[14,15] and *in vivo*[16] through activation of the JAK-STAT pathway and upregulation of interferon-stimulated genes (ISGs)[14,17,18]. The molecular phenotype of this genetic association remains unclear. The GWAS markers have not consistently been associated with hepatic *IFNL3* mRNA expression[19–21], and a nonsynonymous *IFNL3* variant, rs8103142 (encoding a p.Arg70Lys alteration), which is in strong linkage disequilibrium (LD) with rs12979860 in all HapMap populations (r^2 = 0.8–1.0), does not seem to affect the function of the IFNL3 protein[22].

[1]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [2]Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [3]Immunology Section, Liver Diseases Branch, National Institute for Diabetes, Digestive and Kidney Diseases, US National Institutes of Health, Bethesda, Maryland, USA. [4]Division of Therapeutic Proteins, Center for Drug Evaluation & Research, US Food and Drug Administration, Bethesda, Maryland, USA. [5]Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [6]Information Management Services, Silver Spring, Maryland, USA. [7]Department of Epidemiology, The Johns Hopkins School of Hygiene and Public Health, Baltimore, Maryland, USA. [8]Department of Medicine, Carolinas Medical Center, Charlotte, North Carolina, USA. [9]National Development and Research Institutes, New York, New York, USA. [10]Center for the Study of Hepatitis C, Weill Medical College of Cornell University, New York, New York, USA. [11]Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland, USA. [12]Gastroenterology Service, Veterans Affairs (VA) Long Beach Healthcare System, Long Beach, California, USA. [13]Division of Gastroenterology, University of California–Irvine, Irvine, California, USA. [14]Division of Infectious Diseases, Johns Hopkins University, Baltimore, Maryland, USA. Correspondence should be addressed to L.P.-O. (prokuninal@mail.nih.gov) or T.R.O. (obrient@mail.nih.gov).

Received 11 September 2012; accepted 7 December 2012; published online 6 January 2013; doi:10.1038/ng.2521

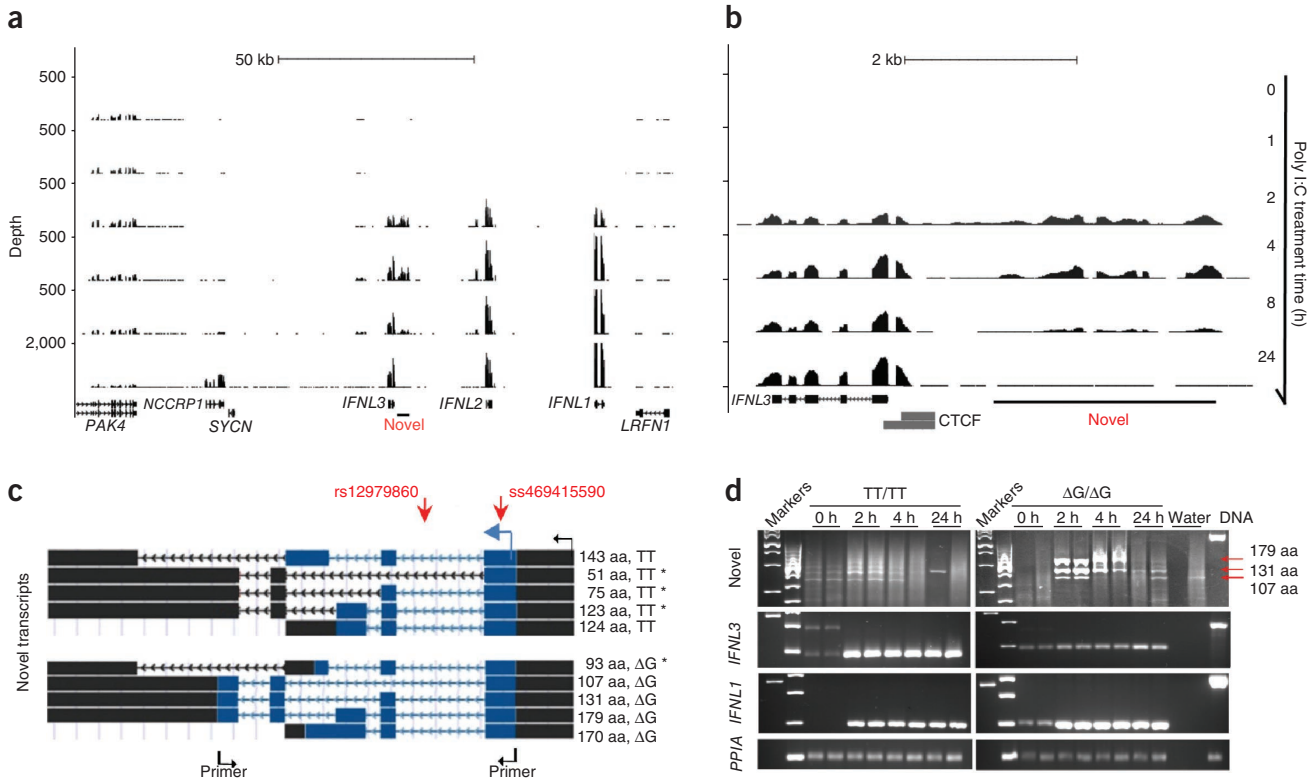

Figure 1 Identification of a novel transcribed region upstream of *IFNL3*. (**a**) RNA-seq in PHHs treated with 50 μg/ml polyI:C for 0, 1, 2, 4, 8 or 24 h. The RNA-seq plot of the 150-kb region in the USCS Genome Browser shows expression of *IFNL1*, *IFNL2*, *IFNL3* and a novel transcribed region upstream of *IFNL3*. The number of reads (depth) corresponds to the level of mRNA expression. (**b**) Detailed view of the RNA-seq results for *IFNL3* and the novel transcribed region. A CTCF transcriptional insulator (Encyclopedia of DNA Elements (ENCODE) data)[24] between these regions indicates their independence. (**c**) Splicing architecture of the ten novel transcripts (NCBI accession numbers are presented in **Supplementary Table 2**). The GWAS marker rs12979860 is located within intron 1, and a novel marker, ss469415590, with TT or ΔG alleles is located within exon 1 and is common to all transcripts. Transcription and translation start sites are marked by black and blue arrows, respectively; ORFs are shaded in blue. Asterisks indicate transcripts that carry premature stop codons and are likely to be eliminated by nonsense-mediated mRNA decay. Arrows indicate the location of the primers (**Supplementary Table 1**) used to generate PCR products. aa, amino acids. (**d**) PCR results in PHH cDNA using the primers from **c**. No distinct PCR product is expected in the sample with the TT/TT genotype. In the sample with the ΔG/ΔG genotype, the primers captured transcripts producing proteins of 179, 131 and 107 amino acids but not of 170 amino acids. A transcript producing a 93-amino-acid protein fragment is expected to be degraded. *IFNL3*, *IFNL1* and *PPIA* (endogenous control) levels were measured in the same samples. Similar amounts of DNase I–treated high-quality RNA were used for all reactions.

RESULTS

A genetic variant creates IFNL4, a novel interferon protein

We performed sequencing of mRNA (RNA-seq) in primary human hepatocytes (PHHs) treated with polyinosinic:polycytidylic acid (polyI:C), which is a synthetic mimic of double-stranded HCV RNA. The PHH sample was from a liver donor who was heterozygous for rs12979860 (C/T) and uninfected with HCV. Hepatocytes were treated with polyI:C for 0, 1, 2, 4, 8 or 24 h, and induction of the IFN-λ genes (*IFNL1*, *IFNL2* and *IFNL3*) was confirmed by TaqMan expression analysis before performing RNA-seq. An analysis of RNA-seq data that focused on a 150-kb region centered on rs12979860 showed concordance with the TaqMan expression results—there was no expression of IFN-λ genes without polyI:C treatment, and these genes were induced after 2–24 h of activation by polyI:C (**Fig. 1a**).

Table 1 Comparisons between p179 (IFNL4) and selected members of the class 2 cytokine family

Protein	p179 (IFNL4)	IFNL3	IFNL1	IFNL2	IFN-α1	IL-10	IL-22	IFN-β1
Type	IFN III?	IFN III	IFN III	IFN III	IFN I	IL-10 family	IL-10 family	IFN I
mRNA	JN806234	NM_172139	NM_172140	NM_172138	NM_002175	NM_000572	NM_020525	NM_002176
Protein ID	AFQ38559	Q8IZI9	Q8IU54	Q45KQ8	P05014	Q6FGW4	Q9GZX6	Q5VWC9
Size, amino acids	179	196	200	200	189	178	179	187
Exons	5	5	5	6	1	5	5	1
Identity with p179, *n* amino acids (%)	–	52 (29.1)	48 (26.8)	47 (26.3)	26 (14.5)	24 (13.4)	20 (11.1)	20 (11.1)
Similarity with p179, *n* amino acids (%)	–	73 (40.8)	66 (36.8)	70 (39.1)	47 (26.2)	40 (22.3)	45 (25.1)	44 (24.5)

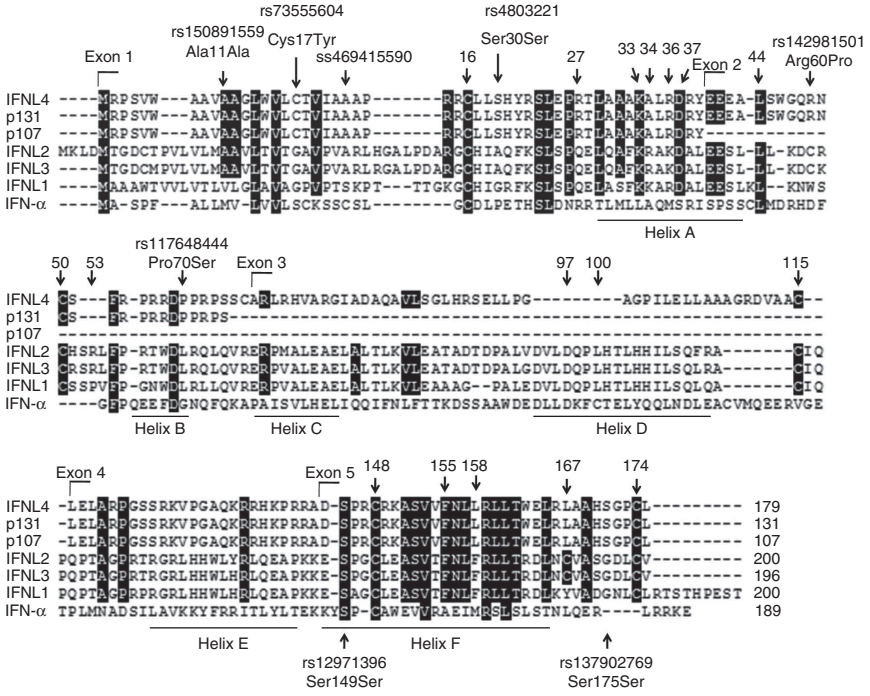

Figure 2 Protein sequence analysis. ClustalW protein sequence alignment for IFNL4 (p179), p131, p107, IFN-λs (IFNL1, IFNL2 and IFNL3) and IFN-α. Identical amino acids are shaded in black. Exon numbering and the location of ss469415590 and other variants identified by sequencing of 270 HapMap samples are based on IFNL4 protein sequence. Annotations of protein helices, amino-acid numbering and specific amino acids are based on the sequence of the mature IFNL3 protein[26,27] (without leader peptide). Cysteines conserved between IFNL4 and all other IFN-λ proteins are indicated (at residues 16, 50, 115, 148 and 174), as are the interaction sites of IFN-λs with their first common receptor IFNLR1 (residues 27, 33, 34, 36, 37, 44, 53, 155 and 158) and with their second receptor, IL10R2 (residues 97 and 100).

We also observed transient activation of a novel transcribed region upstream of *IFNL3*, with the highest levels of expression detected at 2 and 4 h after treatment (**Fig. 1b**). Analysis of paired-end RNA-seq reads identified one major splice junction site. Using this common sequence as a starting point for 5′ rapid amplification of cDNA ends (5′RACE), we mapped a transcription start site, followed by a unique protein translation start site 277 bp downstream. Within exon 1 (at amino acid 22), we detected a novel compound dinucleotide variant, denoted ss469415590 (TT>ΔG), comprised of a one-base insertion or deletion (indel) polymorphism (loss of T, rs67272382) and a one-base substitution variant (T>G, rs74597329). Using polyI:C-stimulated PHHs from five additional liver donors and the primers described in **Supplementary Table 1**, we cloned and annotated ten individual transcripts created by a combination of the ss469415590 alleles and inclusion of several alternative exons (**Fig. 1c**,**d**; NCBI accession numbers are presented in **Supplementary Table 2**). The location of these novel transcripts 3 kb upstream of and in the same orientation as *IFNL3* raised the possibility that they were alternatively spliced forms of *IFNL3* or fusions. However, the presence of a CTCF transcriptional insulator site[23,24] between the two transcribed regions (**Fig. 1b**), the results of the RACE experiments and the inability to generate an RT-PCR product covering *IFNL3* and the novel transcribed region confirmed the independence of these loci. Despite high overall similarity with a genomic region upstream of *IFNL2*, the novel transcripts and ss469415590 are specific for the region upstream of *IFNL3* (**Supplementary Fig. 1**).

Of the ten novel transcripts, four were interrupted by premature stop codons and, thus, are likely to be eliminated by nonsense-mediated mRNA decay[25]. The remaining six transcripts were predicted to produce full-length proteins of 143 amino acids (p143) and 124 amino acids (p124) from transcripts with the ss469415590[TT] allele and 179 amino acids (p179), 170 amino acids (p170), 131 amino acids (p131) and 107 amino acids (p107) from transcripts with the ss469415590[ΔG] allele (**Fig. 1c**). A global protein BLAST search found homology only for p179, with 29.1% amino-acid identity and 40.8% amino-acid similarity with IFNL3. However, the p179 and *IFNL3* cDNA sequences were not similar enough to be aligned using the BLAST bl2seq tool. On the basis of its protein sequence homology with type III IFNs (**Table 1**), the International Society of Interferon and Cytokine Research (ISICR) and the nomenclature committee of the Human Genome Organization (HUGO) designated p179 as interferon λ4 protein (IFN-λ4, IFNL4). IFNL3 and p179 (IFNL4) proteins are most related within the sequences that correspond to the A and F helices of IFNL3, which constitute the core area for the interaction of IFNL3 and other type III IFNs with their primary receptor, IFNLR1 (also known as IL28R1). However, IFNL4 differs in the area corresponding to the D helix of IFNL3, which is the area of interaction of type III IFNs with the second chain of the IFN-λ receptor complex, IL10R2 (**Fig. 2**)[26–29].

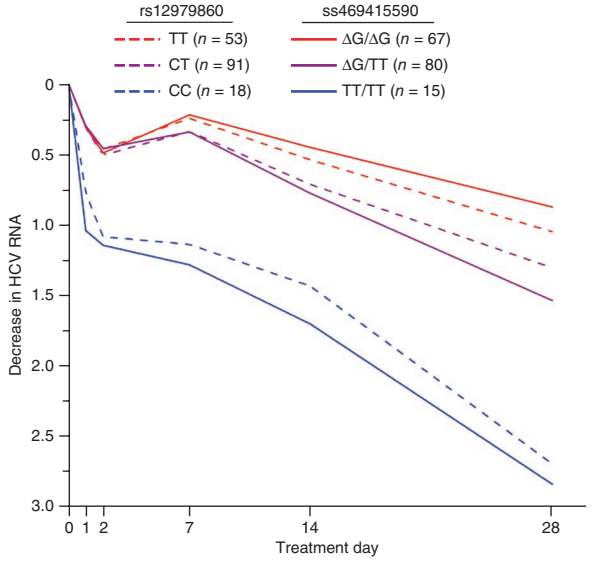

Figure 3 Median decrease in HCV RNA levels ($\log_{10}$ international units (IU)/ml) in African-American participants in the Virahep-C study during the first 28 d of treatment with pegIFN-α/RBV. $P = 0.015$ for the mean differences in HCV RNA levels at day 28 for each of the three genotype groups at ss469415590 relative to the respective rs12979860 genotype groups.

Table 2 Comparison of rs12979860 and ss469415590 genotypes in African-American participants of Virahep-C and HALT-C for associations with the response to treatment with pegIFN-α/RBV

Variant	Genotype	*n*	Treatment time point								
			Week 20/24[a]			End of treatment			SVR		
			Response rate (%)	OR	*P* value	Response rate (%)	OR	*P* value	Response rate (%)	OR	*P* value
Virahep-C (*n* = 169)											
rs12979860	TT	57	40.4	Reference		36.8	Reference		24.6	Reference	
	CT	93	50.5	1.51	0.23	40.9	1.18	0.62	29.0	1.26	0.55
	CC	19	63.2	2.53	0.089	52.6	1.90	0.23	36.8	1.79	0.30
ss469415590	ΔG/ΔG	71	35.2	Reference		32.4	Reference		21.1	Reference	
	ΔG/TT	82	56.1	2.35	0.010	45.1	1.72	0.11	31.7	1.73	0.14
	TT/TT	16	68.8	4.05	0.019	56.3	2.68	0.080	43.8	2.90	0.067
HALT-C Trial (*n* = 144)											
rs12979860	TT	61	9.8	Reference		8.2	Reference		3.3	Reference	
	CT	72	16.7	1.83	0.26	12.5	1.60	0.42	6.9	2.20	0.36
	CC	11	45.5	7.64	6.2×10^{-3}	36.4	6.40	0.018	27.3	11.06	0.015
ss469415590	ΔG/ΔG	68	8.8	Reference		7.4	Reference		2.9	Reference	
	ΔG/TT	68	19.1	2.44	0.090	14.7	2.17	0.18	8.8	3.19	0.16
	TT/TT	8	50.0	10.34	4.7×10^{-3}	37.5	7.56	0.019	25.0	11.0	0.027

[a]In Virahep-C[32], response is at 24 weeks after initiation of pegIFN-α/RBV treatment; in HALT-C[33], response is at 20 weeks after initiation of pegIFN-α/RBV treatment. OR, odds ratio.

Association of genetic variants in *IFNL4* with HCV clearance

The GWAS markers rs12979860 and rs8099917 are located 367 bp downstream (intron 1) and 4 kb upstream of ss469415590, respectively. Analysis of data from the HapMap Project[30] (**Supplementary Table 3**) and the 1000 Genomes Project[31] (**Supplementary Table 4**) showed that the IFNL4-creating ss469415590[ΔG] allele is perfectly correlated with the unfavorable rs12979860[T] allele in Asians (Han Chinese in Beijing, China (CHB) and Japanese in Tokyo, Japan (JPT) sets, $r^2 = 1.00$) and well correlated in Europeans (Utah residents of Northern and Western European ancestry (CEU) set: HapMap, $r^2 = 0.92$; 1000 Genomes Project, $r^2 = 0.83$). In Africans, however, this correlation is only moderate (Yoruba from Ibadan, Nigeria (YRI) set: HapMap, $r^2 = 0.71$; 1000 Genomes Project, $r^2 = 0.65$), even though rs12979860 is the best surrogate for ss469415590 of all markers within the 100-kb region. Correlation between ss469415590 and rs8099917 was high in Asians ($r^2 = 0.91$) and moderate in Europeans ($r^2 = 0.44$) but very low in Africans ($r^2 = 0.008$) (**Supplementary Table 3**).

We assessed the associations of ss469415590 and rs12979860 with HCV clearance in 1,436 African-American and 1,480 European-American individuals from 4 studies. In Virahep-C[32] and HALT-C[33], we evaluated the response to pegIFN-α/RBV therapy in individuals with CHC (**Supplementary Table 5**). There were differences in the rates of sustained virological response (SVR) among the subjects from these studies, which reflect well-known ancestry-related differences in response to treatment and the differing selection criteria for these clinical trials: Virahep-C (European-American), 52%; Virahep-C (African-American), 28%; HALT-C (European-American), 18%; and African-American, 7%. We evaluated spontaneous HCV clearance in injection drug users enrolled in two studies, UHS[34] and ALIVE[35] (**Supplementary Table 6**). The decrease in the amount of HCV RNA during the first 28 d of treatment is a powerful early predictor of ultimate treatment response that is strongly associated with rs12979860 genotype[36,37]. In African-American Virahep-C participants, the decline in HCV RNA levels after 28 d of treatment was more strongly associated with ss469415590 genotype than with rs12979860 genotype ($P = 0.015$, difference in mean values; **Fig. 3** and **Supplementary Table 7**). In the same study, we observed a stronger association for ss469415590 than for rs12979860 with other measures of treatment response (week 24 response, end-of-treatment response and SVR; **Table 2**), although these differences did not reach statistical significance. The association pattern was similar in African-American individuals from the HALT-C study, with a stronger association for ss469415590 than for rs12979860 (week 20 response, end-of-treatment and SVR; **Table 2**). Spontaneous HCV clearance in African-Americans was evaluated using the area under the receiver operating curve (AUC). In UHS participants, the AUC value was greater for ss469415590 (0.62) than for rs12979860 (0.58) (**Table 3**). In the ALIVE study, the AUC values were similar for rs12979860 (0.64) and ss469415590 (0.64) (**Table 3**).

Table 3 Comparison of rs12979860 and ss469415590 genotypes in African-American participants of UHS and ALIVE for associations with the spontaneous clearance of HCV infection

Variant	Genotype	Chronic	Percentage[a]	Clearance	Percentage[a]	OR	*P* value	AUC
UHS (*n* = 459)								
rs12979860	TT	119	34.0	26	23.9	Reference		0.58
	CT	183	52.3	54	49.5	1.35	0.26	
	CC	48	13.7	29	26.6	2.77	1.5×10^{-3}	
ss469415590	ΔG/ΔG	137	39.1	26	23.9	Reference		0.62
	ΔG/TT	168	48	53	48.6	1.66	0.056	
	TT/TT	45	12.9	30	27.5	3.51	7.9×10^{-5}	
ALIVE (*n* = 664)								
rs12979860	TT	228	38.9	15	19.2	Reference		0.64
	CT	275	46.9	37	47.4	2.05	0.025	
	CC	83	14.2	26	33.3	4.76	7.6×10^{-6}	
ss469415590	ΔG/ΔG	255	43.5	19	24.4	Reference		0.64
	ΔG/TT	265	45.2	36	46.2	1.82	0.043	
	TT/TT	66	11.3	23	29.5	4.68	5.5×10^{-6}	

UHS[34]and ALIVE[35] are epidemiological studies of injection drug users who were evaluated for chronic HCV infection on the basis of testing for antibodies to HCV and HCV RNA.
[a]Percentage of subjects with the indicated genotype.

Virahep-C, HALT-C and UHS also enrolled European-American participants. In these subjects, ss469415590 and rs12979860 showed similar associations for both treatment-induced and spontaneous HCV clearance (**Supplementary Tables 7–10**). Taken as a whole, our results show that, in African-American individuals, ss469415590 is a better marker than rs12979860 for predicting response to pegIFN-α/RBV treatment of CHC and possibly for spontaneous HCV clearance, whereas these variants are similarly informative in European-Americans.

By sequencing *IFNL4* in 270 HapMap samples, we annotated 3 nonsynonymous variants, rs73555604 (p.Cys17Tyr) in exon 1 and rs142981501 (p.Arg60Pro) and rs117648444 (p.Pro70Ser) in exon 2, as well as 4 synonymous variants, rs150891559 (p.Ala11Ala) and rs4803221 (p.Ser30Ser) in exon 1 and rs12971396 (p.Ser149Ser) and rs137902769 (p.Ser175Ser) in exon 5 (**Fig. 2**, **Supplementary Fig. 2** and **Supplementary Tables 11** and **12**). On the basis of a haplotype analysis of 16 markers from the 8-kb *IFNL3-IFNL4* region, we identified 8 markers that captured all haplotypes present in the HapMap sets (**Supplementary Table 13**). These eight markers were also tested in European-American and African-American individuals from Virahep-C. In all populations, the unique favorable haplotype included the ss469415590[TT] allele, which eliminates the IFNL4 protein. The unfavorable ss469415590[ΔG] allele was found on a number of haplotypes, including two haplotypes that were reported as being neutral in Europeans despite carrying the unfavorable rs12979860[T] allele[38,39]; these two haplotypes included minor alleles of either of

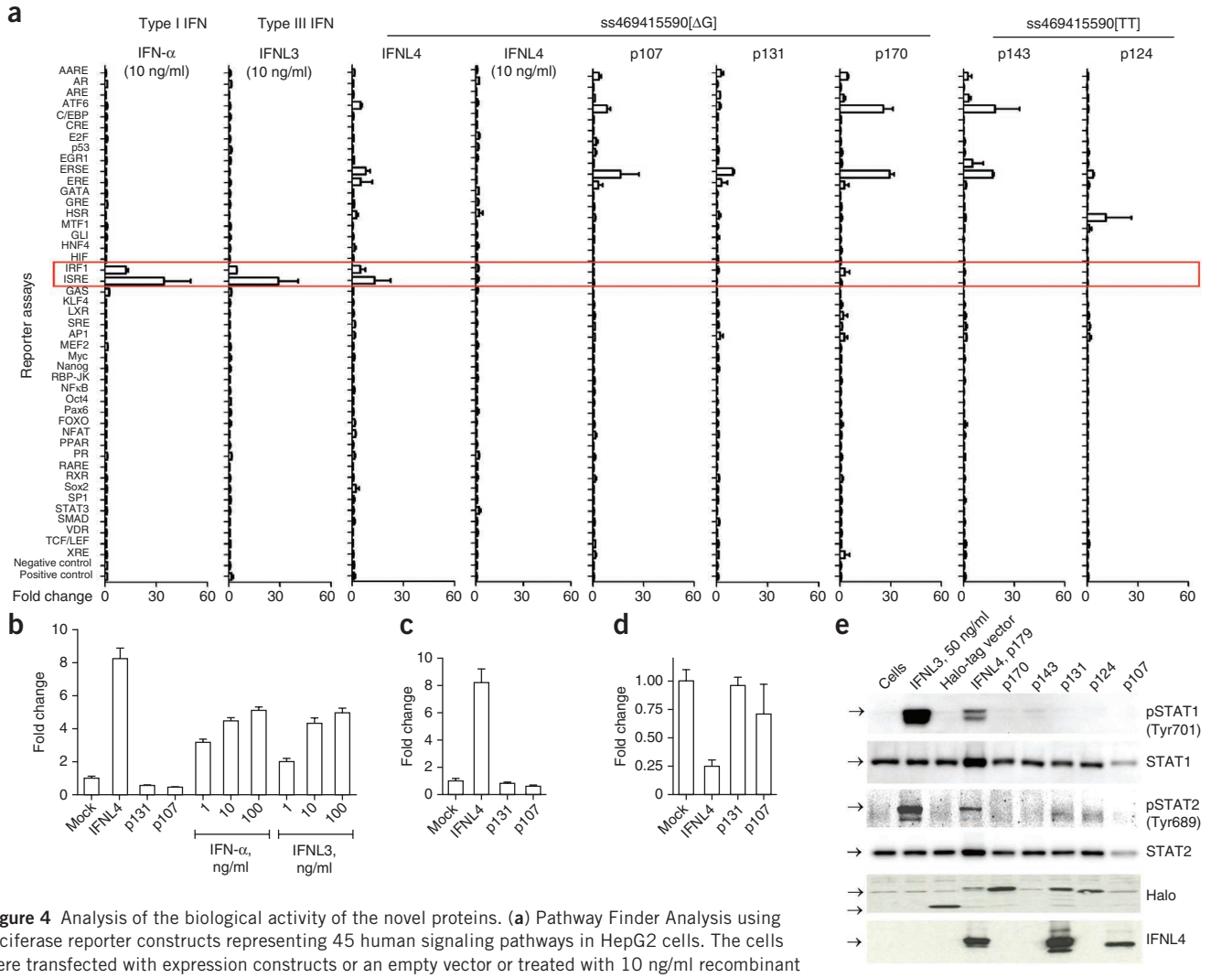

Figure 4 Analysis of the biological activity of the novel proteins. (**a**) Pathway Finder Analysis using luciferase reporter constructs representing 45 human signaling pathways in HepG2 cells. The cells were transfected with expression constructs or an empty vector or treated with 10 ng/ml recombinant purified IFN-α, IFNL3 or IFNL4 or with PBS. All results represent the mean values of two independent biological transfection and/or treatment replicates. Error bars, s.d. The red box indicates reporters significantly induced by treatment with IFN-α, IFNL3 and IFNL4. (**b**) Luciferase activity after transfection with vector expressing IFNL4, p131 or p107 and treatment with recombinant purified IFN-α or IFNL3 in the HepG2 cell line transiently cotransfected with the ISRE-Luc reporter. The results are normalized to the activity seen after transfection with empty vector (mock) and represent the mean values of eight biological replicates. Error bars, s.e.m. (**c**) Luciferase activity after transient transfection with vector expressing IFNL4, p131 or p107 in the HepG2 cell line stably expressing the ISRE-Luc reporter. The results are normalized to the activity seen after transfection with empty vector (mock) and represent the mean values of 11 biological replicates. Error bars, s.e.m. (**d**) Test for antiviral effects of the expression constructs for IFNL4, p131 and p107 transiently transfected into Huh7-Lunet cells stably expressing a subgenomic luciferase-expressing HCV replicon (HCV-Luc) compared to the effect seen after transfection with empty vector (mock). Results represent the mean values of four biological replicates. Error bars, s.e.m. (**e**) Protein blot analysis of STAT1 phosphorylated at Tyr701 (pSTAT1) and STAT2 phosphorylated at Tyr689 (pSTAT2) in HepG2 cells transiently transfected with constructs expressing the six protein isoforms. All constructs encoding the Halo tag produce proteins detectable with an antibody for Halo; the rabbit monoclonal antibody to IFNL4 recognizes p179 as well as the nonfunctional isoforms p131 and p107.

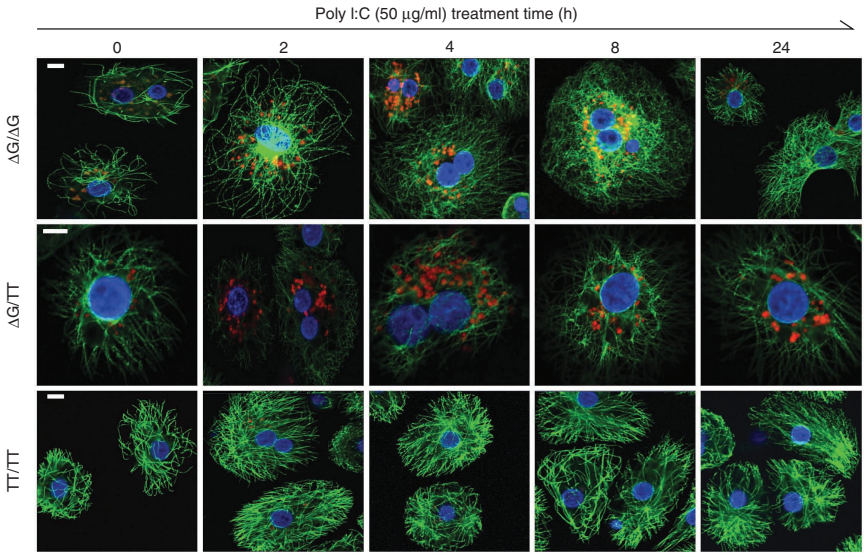

Figure 5 Confocal imaging of IFNL4 in PHHs from liver donors with different ss469415590 genotypes. Confocal imaging in PHHs treated with 50 μg/ml of polyI:C for 0, 2, 4, 8 or 24 h. Red, IFNL4; green, cytoskeleton (α-tubulin), blue, nuclei. Confocal imaging of endogenous IFNL4 expression in the same ss469415590 samples with ΔG/TT genotype after *in vitro* infection with the HCV JFH1 strain is presented in **Supplementary Figure 6**. Scale bars, 10 μm.

the rs73555604 and rs11764844 nonsynonymous variants. It is possible, therefore, that these variants modify the risk in carriers of the unfavorable ss469415590[ΔG] allele and are the source of the haplotype heterogeneity previously reported in Europeans[38,39]; however, data from Virahep-C are too sparse to confirm this theory (**Supplementary Table 13**).

IFNL4 induces expression of ISGs

We evaluated the functional properties of the six novel protein isoforms created by alleles of ss469415590. For an analysis of 45 signaling pathways, HepG2 hepatoma cells were transiently transfected with expression constructs for all six isoforms or treated with recombinant IFN-α, IFNL3 or IFNL4 (with the latter two produced in the sfs9 baculoviral expression system). Only transfection with construct expressing IFNL4, as well as treatment with IFN-α or IFNL3, induced activation of an interferon-stimulated response element luciferase reporter (ISRE-Luc), which contains STAT1- and STAT2-binding sites responsive to type I and type III IFN signaling and the IRF1 reporter (**Fig. 4a**). These results were validated in HepG2 cells transiently (**Fig. 4b**), as well as stably (**Fig. 4c**), expressing ISRE-Luc reporter constructs. The effect was comparable when the cells were transfected with constructs expressing IFNL4 that generated proteins with either a Halo or a Flag tag (**Supplementary Fig. 3**). Similarly, only transient transfection with the construct expressing IFNL4 decreased HCV RNA replication in hepatoma cells stably expressing a subgenomic luciferase-expressing hepatitis C virus replicon[40] (**Fig. 4d**) and induced STAT1 and STAT2 phosphorylation (**Fig. 4e**). Transfection with IFNL4 activated the ISRE reporter in HepG2 and 293T cells but not in HeLa cells (**Supplementary Fig. 4**).

Recombinant IFNL4 protein expression was detectable in the cells and cell lysates of transfected HepG2 and 293T cells by confocal imaging and protein blots, respectively, with antibodies specific for IFNL4 and tag proteins (**Supplementary Figs. 5** and **6**). We also detected weak expression of recombinant IFNL4 in the medium of transfected HepG2 cells but not 293T cells (**Supplementary Fig. 7**). In polyI:C-stimulated PHHs from liver donors not infected with HCV, the endogenous expression of IFNL4 protein was detected by confocal imaging in carriers of the unfavorable ss469415590[ΔG] allele but not in a homozygous carrier of the favorable ss469415590[TT] allele (**Fig. 5**). In hepatocytes from the donor heterozygous for ss469415590[ΔG], we detected endogenous expression of IFNL4 in cells treated with polyI:C and after *in vitro* infection with the JFH1 HCV strain (**Supplementary Fig. 6d**). In fact, in hepatocytes from one of these donors (ΔG/TT genotype), we observed low IFNL4 expression, even without polyI:C treatment or *in vitro* HCV infection (at the time point at 0 h). Although preliminary, these results suggest that IFNL4 might be expressed in conditions unrelated to HCV infection.

To further explore the functional consequences of IFNL4 expression, we performed RNA-seq in HepG2 cells transiently transfected with empty vector or construct expressing IFNL4 (**Supplementary Fig. 8a**) and found that the top canonical pathways induced by IFNL4 were related to the activation of interferon signaling (**Supplementary Fig. 8b**). We validated the RNA-seq results by quantitative RT-PCR (qRT-PCR) analysis and showed that IFNL4 induced the expression of many ISGs in a pattern similar to that induced by IFN-α and IFNL3 (**Supplementary Fig. 8c**). Previously, many of these ISGs have been shown to be expressed at higher levels in liver biopsies taken before treatment from HCV-infected patients who do not respond to pegIFN-α/RBV treatment; these individuals tend to carry the unfavorable genotypes at rs12979860 and rs8099917 (ref. 19,41–44), which mark the ss469415590[ΔG] allele that produces IFNL4. To mimic this clinical phenotype, we transfected HepG2 cells with empty vector or IFNL4 expression constructs and/or treated cells with 10 ng/ml of recombinant IFN-α or IFNL3. In these samples, we validated the RNA-seq data by qRT-PCR analysis and showed that IFNL4 induced expression of selected ISGs (*STAT1*, *ISG15*, *IFIHI1-MDA5*, *OAS1*, *MX1* and *DHX58-RIG-I*) in a pattern similar to that induced by IFN-α and IFNL3 (**Supplementary Fig. 8d**). Furthermore, treatment with IFN-α or IFNL3 of cells already expressing IFNL4 did not induce additional activation of ISGs (**Supplementary Fig. 8d**). Some genes known as markers of HCV-induced liver damage, such as the chemokine *CCL5* (also known as *RANTES*)[45] and the proto-oncogene *FOS*[46–48], were induced by IFNL4 but not by IFNs (**Supplementary Fig. 8d**), suggesting a functional role for IFNL4 distinct from the roles of IFN-α and IFNL3.

DISCUSSION

We have identified a new inducible human protein-coding gene, *IFNL4*, which is related to but distinct from known IFNs and other class 2 cytokines. The 179 amino-acid ORF of the *IFNL4* transcript is created by a common deletion frameshift allele of ss469415590, which is a dinucleotide variant strongly linked with rs12979860. In individuals of African ancestry, the IFNL4-generating ss469415590[ΔG] allele is superior to the rs12979860[T] allele in predicting poorer response to pegIFN-α/RBV treatment of CHC. Within *IFNL4*, we identified three nonsynonymous variants, rs73555604 (p.Cys17Tyr), rs142981501 (p.Arg60Pro) and rs11764844 (p.Pro70Ser), which are present on haplotypes with the ss469415590[ΔG] allele. The effects of these variants

on IFNL4 biological function and their impact on HCV clearance in different populations should be further explored.

Analysis of the genomic sequences of 45 species for which the sequence of the *IFNL4* region is available in the UCSC Genome Browser showed that the unfavorable IFNL4-generating ss469415590[ΔG] allele is an ancestral variant present in all species. The existence of IFNL4 protein could be predicted only in the genomes of macaques (marmosets and rhesus), orangutans, chimpanzees and humans (**Supplementary Fig. 9**). The beneficial insertion ss469415590[TT] allele seems to be a recently derived variant, which became common in all human populations (93% frequency in Asians, 68% frequency in Europeans and 23% frequency in Africans, according to HapMap samples), suggesting positive selection for this allele. The introduction of frameshifts is considered to be an evolutionary mechanism for the rapid emergence of new proteins[49,50], but, in this case, an insertion allele that abrogates IFNL4 seems to have been selected during evolution.

We found that the IFNL4 protein of 179 amino acids induces STAT1 and STAT2 phosphorylation, activates the ISRE-Luc reporter and ISGs, and generates antiviral response in hepatoma cells. The mechanisms by which IFNL4 induces these responses but nevertheless impairs HCV clearance are currently under investigation. IFNL4 and IFNL3 have similar residues in the area that is known to interact with the primary receptor of IFNL3 (IFNLR1) but differ in the region of IFNL3 that interacts with the second chain of the IFNL receptor complex, IL10R2. Thus, it is possible that IFNL4 activates JAK-STAT signaling through a unique receptor complex consisting of IFNLR1 and a currently undefined second receptor chain or that IFNL4 functions as a decoy cytokine competing with type III IFNs for the binding of IFNLR1. We also found that the IFNL4-caused preactivation of interferon signaling prevents further activation by type I and type III IFNs. We used an allele-specific mRNA expression assay (**Supplementary Fig. 10**) and explored endogenous *IFNL4* expression in PHHs, where it was induced by polyI:C, IFN-α (but not IFNL3) treatment and *in vitro* infection with HCV (**Supplementary Fig. 11**). However, no *IFNL4* mRNA expression was induced by polyI:C, IFN-α or IFNL3 in several transformed cell lines that carry the ss469415590[ΔG] allele (HepG2, HeLa, 293T, A549 and HH29). Experiments aimed at elucidating the triggers of IFNL4 expression in diverse conditions and cell types and its receptor components are ongoing and may provide greater insight regarding its mechanism of action.

Previous studies found that individuals with CHC who carry rs12979860[T], which marks the ss469415590[ΔG] allele, before treatment have somewhat higher hepatic expression of ISGs but poorer ISG response to pegIFN-α/RBV treatment[19,41–44]. The rs12979860[T] variant has also been associated with lower HCV RNA levels in the absence of treatment[9,51]. Our *in vitro* experiments in HepG2 cells showed that IFNL4 induced activation of ISGs, which was not further enhanced by exogenous treatment with IFN-α or IFNL3. Taken together, these data suggest that IFNL4 induces weak expression of ISGs that provides an antiviral response that might lower the HCV load, although it also reduces the responsiveness to type I and type III IFNs that is needed for efficient HCV clearance.

It has been reported that rs12979860 predicts early viral kinetics in HCV-infected patients receiving IFN-α–free treatment[52]. This genotype has also been associated with the response to IFN-α–based treatment of chronic hepatitis B virus (HBV) infection in some studies[53,54]. Furthermore, IFN-α therapy is used for a number of other clinical conditions, including some forms of cancer[55]. Thus, therapeutic inhibition of IFNL4 might represent a novel biological strategy for the treatment of HCV and HBV infection and possibly other diseases, and *IFNL4* genotype could be used to select patients for this therapy.

URLs. UCSC Genome Browser, http://genome.ucsc.edu/; IGV, http://www.broadinstitute.org/igv/; full list of the reporter constructs tested, http://www.sabiosciences.com/reporter_assay_product/HTML/CCA-901L.html.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Sequences for all novel transcripts JN806234, JN806225, JN806226, JN806232, JN806233, JN806227, JN806228, JN806229, JN806230 and JN806231 are deposited at NCBI. ss469415590 and ss539198934 have been submitted to the dbSNP database.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the National Cancer Institute Sequencing Core Facility for help with RNA-seq. This research was supported by the Intramural Research Program of the US National Institutes of Health (National Cancer Institute, Division of Cancer Epidemiology and Genetics; National Institute for Diabetes, Digestive and Kidney Diseases), as well as the following grants: DA R01 013324 (J.A. and D.L.T.); ALIVE cohort, R01-DA-04334 and R01-DA-12568, and US National Institutes of Health grants R01-DA09532, R01-DA12109, R01-DA13245 and R01-DA16159 (B.R.E.); National Cancer Institute contracts N02-CP-91027 and N01-CO-12400 (B.R.E.); Substance Abuse and Mental Health Services Administration grant H79-TI12103 (B.R.E.); and the National Cancer Institute Director's Innovation Award (L.P.-O.). The Virahep-C and HALT-C studies were conducted, respectively, by the Virahep-C and HALT-C Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases. The data and samples from the Virahep-C and HALT-C studies reported here were supplied by the National Institute of Diabetes and Digestive and Kidney Diseases Central Repositories. This manuscript was not prepared in collaboration with the Virahep-C study group or the HALT-C study group and does not necessarily reflect the opinions or views of the Virahep-C Trial and HALT-C Trial, the National Institute of Diabetes and Digestive and Kidney Diseases Central Repositories or the National Institute of Diabetes and Digestive and Kidney Diseases. The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services nor does mention of trade names, commercial products or organizations imply endorsement by the US government.

AUTHOR CONTRIBUTIONS

L.P.-O. and T.R.O. conceived and supervised the project. J.A., H.L.B., B.R.E., C.D.H., T.R.M. and D.L.T. performed the clinical and epidemiological studies from which DNA samples and data were collected. L.P.-O., B.R. and R.P.D. designed the experiments. L.P.-O., B.M., W.T., H.P., H.D., D.H., P.P.-G., I.K., A.M., N.B., M.T., L.L., F.S., B.R. and R.P.D. performed experiments and analysis. T.R.O., R.M.P. and S.C. performed statistical analysis of genetic association. L.P.-O. and T.R.O. wrote the manuscript. All authors contributed to scientific discussions and approved the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at http://www.nature.com/doifinder/10.1038/ng.2521.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. NIH. National Institutes of Health Consensus Development Conference Statement. Management of hepatitis C 2002 (June 10–12, 2002). *Gastroenterology* **123**, 2082–2099 (2002).
2. NIH. NIH Consensus Statement on Management of Hepatitis C: 2002. *NIH Consens. State Sci. Statements* **19**, 1–46 (2002).
3. El-Serag, H.B. Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology* **142**, 1264–1273 e1 (2012).
4. European Association of the Study of the Liver. 2011 European Association of the Study of the Liver hepatitis C virus clinical practice guidelines. *Liver Int.* **32** (suppl. 1), 2–8 (2012).
5. Sarrazin, C., Hezode, C., Zeuzem, S. & Pawlotsky, J.M. Antiviral strategies in hepatitis C virus infection. *J. Hepatol.* **56** (suppl. 1), S88–S100 (2012).

6. Ciesek, S. & Manns, M.P. Hepatitis in 2010: the dawn of a new era in HCV therapy. *Nat. Rev. Gastroenterol. Hepatol.* **8**, 69–71 (2011).
7. Rauch, A. *et al.* Genetic variation in *IL28B* is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* **138**, 1338–1345 (2010).
8. Thomas, D.L. *et al.* Genetic variation in *IL28B* and spontaneous clearance of hepatitis C virus. *Nature* **461**, 798–801 (2009).
9. Ge, D. *et al.* Genetic variation in *IL28B* predicts hepatitis C treatment–induced viral clearance. *Nature* **461**, 399–401 (2009).
10. Suppiah, V. *et al. IL28B* is associated with response to chronic hepatitis C interferon-α and ribavirin therapy. *Nat. Genet.* **41**, 1100–1104 (2009).
11. Tanaka, Y. *et al.* Genome-wide association of *IL28B* with response to pegylated interferon-α and ribavirin therapy for chronic hepatitis C. *Nat. Genet.* **41**, 1105–1109 (2009).
12. Kotenko, S.V. *et al.* IFN-λs mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat. Immunol.* **4**, 69–77 (2003).
13. Sheppard, P. *et al.* IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat. Immunol.* **4**, 63–68 (2003).
14. Marcello, T. *et al.* Interferons α and λ inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology* **131**, 1887–1898 (2006).
15. Robek, M.D., Boyd, B.S. & Chisari, F.V. λ interferon inhibits hepatitis B and C virus replication. *J. Virol.* **79**, 3851–3854 (2005).
16. Muir, A.J. *et al.* Phase 1b study of pegylated interferon λ1 with or without ribavirin in patients with chronic genotype 1 hepatitis C virus infection. *Hepatology* **52**, 822–832 (2010).
17. Kempuraj, D. *et al.* Interleukin-28 and 29 (IL-28 and IL-29): new cytokines with anti-viral activities. *Int. J. Immunopathol. Pharmacol.* **17**, 103–106 (2004).
18. Zhou, Z. *et al.* Type III interferon (IFN) induces a type I IFN–like response in a restricted subset of cells through signaling pathways involving both the Jak-STAT pathway and the mitogen-activated protein kinases. *J. Virol.* **81**, 7749–7758 (2007).
19. Honda, M. *et al.* Hepatic ISG expression is associated with genetic variation in interleukin 28B and the outcome of IFN therapy for chronic hepatitis C. *Gastroenterology* **139**, 499–509 (2010).
20. Marukian, S. *et al.* Hepatitis C virus induces interferon-λ and interferon-stimulated genes in primary liver cultures. *Hepatology* **54**, 1913–1923 (2011).
21. Dill, M.T. *et al.* Interferon-induced gene expression is a stronger predictor of treatment response than *IL28B* genotype in patients with hepatitis C. *Gastroenterology* **140**, 1021–1031 (2011).
22. Urban, T.J. *et al. IL28B* genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. *Hepatology* **52**, 1888–1896 (2010).
23. Phillips, J.E. & Corces, V.G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211 (2009).
24. Rosenbloom, K.R. *et al.* ENCODE whole-genome data in the UCSC Genome Browser: update 2012. *Nucleic Acids Res.* **40**, D912–D917 (2012).
25. Shyu, A.B., Wilkinson, M.F. & van Hoof, A. Messenger RNA regulation: to translate or to degrade. *EMBO J.* **27**, 471–481 (2008).
26. Trivella, D.B., Ferreira-Junior, J.R., Dumoutier, L., Renauld, J.C. & Polikarpov, I. Structure and function of interleukin-22 and other members of the interleukin-10 family. *Cell Mol. Life Sci.* **67**, 2909–2935 (2010).
27. Gad, H.H. *et al.* Interferon-λ is functionally an interferon but structurally related to the interleukin-10 family. *J. Biol. Chem.* **284**, 20869–20875 (2009).
28. Yoon, S.I., Logsdon, N.J., Sheikh, F., Donnelly, R.P. & Walter, M.R. Conformational changes mediate interleukin-10 receptor 2 (IL-10R2) binding to IL-10 and assembly of the signaling complex. *J. Biol. Chem.* **281**, 35088–35096 (2006).
29. Yoon, S.I. *et al.* Structure and mechanism of receptor sharing by the IL-10R2 common chain. *Structure* **18**, 638–648 (2010).
30. International HapMap Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).
31. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
32. Conjeevaram, H.S. *et al.* Peginterferon and ribavirin treatment in African American and Caucasian American patients with hepatitis C genotype 1. *Gastroenterology* **131**, 470–477 (2006).
33. Di Bisceglie, A.M. *et al.* Prolonged therapy of advanced chronic hepatitis C with low-dose peginterferon. *N. Engl. J. Med.* **359**, 2429–2441 (2008).
34. Shebl, F.M. *et al. IL28B* rs12979860 genotype and spontaneous clearance of hepatitis C virus in a multi-ethnic cohort of injection drug users: evidence for a supra-additive association. *J. Infect. Dis.* **204**, 1843–1847 (2011).
35. Vlahov, D. *et al.* The ALIVE study, a longitudinal study of HIV-1 infection in intravenous drug users: description of methods and characteristics of participants. *NIDA Res. Monogr.* **109**, 75–100 (1991).
36. Thompson, A.J. *et al.* Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* **139**, 120–129 e18 (2010).
37. Howell, C.D. *et al.* Single nucleotide polymorphism upstream of interleukin 28B associated with phase 1 and phase 2 of early viral kinetics in patients infected with HCV genotype 1. *J. Hepatol.* **56**, 557–563 (2012).
38. Smith, K.R. *et al.* Identification of improved *IL28B* SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort. *Genome Med.* **3**, 57 (2011).
39. Fischer, J. *et al.* Combined effects of different interleukin-28B gene variants on the outcome of dual combination therapy in chronic hepatitis C virus type 1 infection. *Hepatology* **55**, 1700–1710 (2012).
40. Jo, J. *et al.* Analysis of $CD8^+$ T-cell-mediated inhibition of hepatitis C virus replication using a novel immunological model. *Gastroenterology* **136**, 1391–1401 (2009).
41. Dill, M.T. *et al.* Interferon-induced gene expression is a stronger predictor of treatment response than *IL28B* genotype in patients with hepatitis C. *Gastroenterology* **140**, 1021–1031 (2011).
42. Asahina, Y. *et al.* Association of gene expression involving innate immunity and genetic variation in interleukin 28B with antiviral response. *Hepatology* **55**, 20–29 (2012).
43. Onomoto, K. *et al.* Dysregulation of IFN system can lead to poor response to pegylated interferon and ribavirin therapy in chronic hepatitis C. *PLoS ONE* **6**, e19799 (2011).
44. Sarasin-Filipowicz, M. *et al.* Interferon signaling and treatment outcome in chronic hepatitis C. *Proc. Natl. Acad. Sci. USA* **105**, 7034–7039 (2008).
45. Katsounas, A., Schlaak, J.F. & Lempicki, R.A. CCL5: a double-edged sword in host defense against the hepatitis C virus. *Int. Rev. Immunol.* **30**, 366–378 (2011).
46. Tappero, G. *et al.* Intrahepatic expression of *c-fos*, *c-myb* and *c-myc* oncogenes: correlation with virus-induced chronic liver disease and response to interferon. *Ital. J. Gastroenterol. Hepatol.* **29**, 148–154 (1997).
47. Smeyne, R.J. *et al.* Continuous *c-fos* expression precedes programmed cell death in vivo. *Nature* **363**, 166–169 (1993).
48. Kang, S.M. *et al.* c-Fos regulates hepatitis C virus propagation. *FEBS Lett.* **585**, 3236–3244 (2011).
49. Okamura, K., Feuk, L., Marques-Bonet, T., Navarro, A. & Scherer, S.W. Frequent appearance of novel protein-coding sequences by frameshift translation. *Genomics* **88**, 690–697 (2006).
50. White, S.H. The evolution of proteins from random amino acid sequences: II. Evidence from the statistical distributions of the lengths of modern protein sequences. *J. Mol. Evol.* **38**, 383–394 (1994).
51. Uccellini, L. *et al.* HCV RNA levels in a multiethnic cohort of injection drug users: human genetic, viral and demographic associations. *Hepatology* **56**, 86–94 (2012).
52. Chu, T.W. *et al.* Effect of *IL28B* genotype on early viral kinetics during interferon-free treatment of patients with chronic hepatitis C. *Gastroenterology* **142**, 790–795 (2012).
53. Sonneveld, M.J. *et al.* Polymorphisms near *IL28B* and serologic response to peginterferon in HBeAg-positive patients with chronic hepatitis B. *Gastroenterology* **142**, 513–520 e1 (2012).
54. Lampertico, P. *et al. IL28B* polymorphisms predict interferon-related HBsAg seroclearance in genotype D HBeAg-negative patients with chronic hepatitis B. *Hepatology* published online; doi:10.1002/hep.25749 (2 April 2012).
55. Davar, D., Tarhini, A.A. & Kirkwood, J.M. Adjuvant therapy for melanoma. *Cancer J.* **18**, 192–202 (2012).

ONLINE METHODS

Cells. Fresh PHHs from HCV-free liver donors were purchased from Lonza or Celsis. PHHs received in suspension were cultured overnight in collagen-coated plates or chamber slides (BD Biosciences), and unattached PHHs were removed before treatment. Hepatoma HepG2 cells, cervical carcinoma HeLa cells and embryonic kidney 293T cells (all from American Type Culture Collection (ATCC)) were cultured in DMEM supplemented with 10% FBS. Hepatoma Huh7-Lunet cells harboring a subgenomic luciferase-expressing HCV JFH1 replicon were cultured as previously described[40]. PHHs were treated with 50 μg/ml polyI:C (Imgenex) or with 100 ng/ml IFN-α (PBL Interferon Source) or IFNL3 (custom). Cell lines were treated with 10 ng/ml IFN-α, IFNL3 or IFNL4 (custom).

Transfections. Transfections were performed with Lipofectamine LTX Reagent and Opti-MEM medium using standard protocols (Life Technologies).

HCV infection of PHHs. PHHs (0.35×10^5 per well) were infected with HCV (JFH1 strain)[56] for 0, 6 or 24 h in collagen-coated 24-well plates or chamber slides at a multiplicity of infection (MOI) of 2.

Antiviral assays. Huh7-Lunet cells were transfected with constructs expressing IFNL4, p131 or p107 or with empty GFP vector in a 48-well plate, and luciferase expression was measured 48 h after transfection.

Extraction of DNA, RNA and protein. DNA and RNA were prepared using a DNeasy or RNeasy kit with DNase I treatment (Qiagen), respectively, and evaluated by NanoDrop 8000 (Thermo Scientific) and Bioanalyzer 2100 (Agilent Technologies). Protein was prepared by lysing cells in RIPA buffer with protease inhibitors (Sigma-Aldrich). Cell medium was concentrated by 10× and 100× using 9K MWCO protein concentrator tubes (Thermo Scientific).

Protein blotting. Proteins were resolved by 4–12% Tris-glycine SDS-PAGE (Life Technologies). Detection was carried out using the custom mouse and rabbit monoclonal antibodies to IFNL4, rabbit antibody to Halo (Promega) and secondary horseradish peroxidase (HRP)-conjugated goat antibodies to rabbit (sc-2030) or mouse (sc-2031) (Santa Cruz Biotechnology). Signals were detected with the ECL Plus Western blotting detection system (GE Healthcare Life Sciences).

Analysis of STAT1 and STAT2 phosphorylation. HepG2 cells in 6-well plates were untreated, transfected with expression constructs or empty Halo-tag vector or treated for1 h at 37 °C with 50 ng/ml recombinant IFNL3. Equal amounts (50 μg per lane) of whole-cell lysates prepared 48 h after transfection were used for analysis by protein blotting. Detection was performed as previously described[57] with rabbit antibody to phosphorylated Tyr701 of STAT1 (Cell Signaling Technology, 9171) and rabbit antibody to phosphorylated Tyr689 of STAT2 (Millipore, 07-224). Blots were stripped and reprobed with rabbit antibodies to STAT1 or STAT2 (Santa Cruz Biotechnology, sc-417 and sc-476, respectively) to measure the amounts of total protein.

RNA-seq. Total RNA (1 μg) from PHHs or HepG2 cells was used for library preparation with the TruSeq PolyA kit (Illumina). Sequencing by Genome Analyzer (GAIIx) generated 47 million 107-bp paired-end sequencing reads per sample. The TopHat v1.2.0 settings were changed to allow multiple read alignments (up to ten regions, UCSC hg19) and three nucleotide mismatches per each 25-bp segment. Results were viewed with the UCSC Genome Browser and the Integrative Genomics Viewer (IGV) software (see URLs).

Identification and cloning of novel spliced forms. Rapid amplification of cDNA ends (with 5′ and 3′RACE) and cloning of full-length ORFs were performed with the SMARTer RACE cDNA kit (Clontech,) using a pool of DNase I–treated RNA samples from polyI:C-treated PHHs from several liver donors (primers sequences are given in **Supplementary Table 1**). PCR reactions were performed with AmpliTaq Gold 360 Master Mix (Life Technologies) and 360 GC Enhancer (Life Technologies) using a touchdown PCR program consisting of 10 min at 95 °C, 20 cycles of touchdown (30 s at 95 °C, 45 s starting at 70 °C and decreasing by 1 °C every two cycles to 60 °C, 45 s at 72 °C), 25 additional cycles (30 s at 95 °C, 45 s at 60 °C, 45 s at 72 °C) and a final extension time of 7 min at 72 °C. Gel-purified PCR fragments were cloned into the pFC14A expression vector encoding a C-terminal Halo tag (Promega) and sequenced for validation. The IFNL3-Halo expression construct was generated using the same approach. p179 was cloned into a the pcDNA3.1 expression vector encoding a Flag tag.

Production of recombinant proteins. IFNL4 and IFNL3 bacmids were generated in the pFastBac vector encoding a C-terminal His tag (Life Technologies) and expressed in a sf9 baculoviral strain. Using antibody to His tag (Sigma, H1029), expression of IFNL3 was detected in cell medium, which was used for protein purification. Expression of p179 was not detectable in cell medium, and protein was purified from the cell pellet after the cultivation of cells for 3–5 d in 2 l of SF-900 III medium (Life Technologies). Proteins were first purified on the HisTrap 5-ml nickel column and then on the size-exclusion chromatography preparative TSK G3000pw column (21.5 × 600 mm; Tosoh). Purified protein fractions were concentrated and dialyzed into a buffer (20 mM HEPES buffer, 600 mM NaCl,1 mM DTT, 1 mM EDTA, 0.1% NP-40, 10% glycerol). High protein purity (>90%) was confirmed by Coomassie staining and protein blot analyses with antibody to His (H1029, Sigma, 1:3,000 dilution), antibody to IFNL3 and custom mouse and rabbit monoclonal antibodies to p179. The IFNL3 and p179 proteins were custom produced by Protein One.

Development of antibodies to IFNL4. A mouse monoclonal antibody was custom developed for a synthetic peptide KALRDRYEEEALSWGQRNCSF RPRRDSPRPS corresponding to amino acids 44–74 of the p179 protein by Precision Antibody. A rabbit monoclonal antibody was custom developed for a synthetic peptide PGSSRKVPGAQKRRHKPRRADSPRC corresponding to amino acids 128–152 of the p179 protein by Epitomics.

Evaluation of the biological activity of the novel proteins. Luciferase Cignal 45–Pathway Finder Reporter Arrays were used in HepG2 cells according to the manufacturer's instructions (Qiagen). Cells were transfected with expression constructs for p179, p170, p143, p131, p124 and p107 or treated for 24 h with purified recombinant proteins, including 10 ng/ml of IFN-α or IFNL3 and/or IFNL4. Validation was performed with an individual ISRE-Luc Cignal reporter. All studies were performed using at least eight biological replicates. A HepG2 cell line stably expressing the same ISRE-Luc reporter construct was generated by transducing cells with the Luciferase Cignal Lenti ISRE reporter construct (Qiagen) and selecting for positive clones by growth in DMEM supplemented with 10% FBS, 1× antibiotic-antimycotic (Life Technologies) and 2 μg/ml puromycin. The best HepG2 clones with stable integration of the ISRE-Luc reporter were identified through testing with purified recombinant IFN-α and IFNL3.

Global analysis of the transcriptome and pathway analysis. HepG2 cells were transfected with empty Halo tag vector or with the vector expressing IFNL4-Halo. High-quality RNA (RNA integrity number (RIN) of ~10) prepared from transfected cells (after 48 h) was used for sequencing with the HiSeq 2000 (Illumina), generating ~300 million reads per sample. Standard analysis identified 535 transcripts with >2-fold difference in expression and a false discovery rate of <0.05. Ingenuity Pathway Analysis (IPA) performed on this set nominated a list of pathways and specific transcripts. mRNA expression of selected transcripts was evaluated in samples transfected with empty vector or with vector expressing IFNL4, p107 or p131 and/or treated with 10 ng/ml IFN-α or IFNL3 in four biological replicates. mRNA expression in all samples was evaluated with pathway-based RT^2 Profiler PCR arrays according to the manufacturer's instructions (Qiagen).

Confocal imaging. PHHs cultured on collagen-coated slides were treated with 50 μg/ml polyI:C for 0 (untreated), 2, 4, 8 or 24 h. HepG2 cells were transiently transfected with vector expressing IFNL4-Halo. Cells were fixed for 20 min with 4% formaldehyde (Sigma) in PBS, permeabilized with 0.5% Triton X-100 (Sigma) for 5 min, blocked with 4% BSA (Sigma) and incubated overnight at 4 °C with primary antibodies, including mouse monoclonal antibody to IFNL4 (custom), rabbit antibody to α-tubulin (ab-15246, Abcam), rabbit antibody to Halo (Promega), mouse antibody to α-tubulin (ab-7291, Abcam) (all at 1:1,000 dilution). Secondary antibodies included Alexa Fluor 594– and

Alexa Fluor 488–conjugated donkey antibodies to rabbit or mouse, all used at a 1:1,000 dilution (Life Technologies). Slides were covered with mounting medium (Prolong Gold Antifade Reagent with DAPI). Immmunofluorescence images were obtained with a confocal laser-scanning microscope (LSM 510 META, Carl Zeiss).

Analysis of *IFNL4* mRNA expression. All expression assays were purchased from Life Technologies (**Supplementary Fig. 10** and **Supplementary Table 1**). Expression analysis was performed with gene expression master mix (Life Technologies) on DNase I–treated RNA samples on an ABI SDS 7700 instrument. Expression was measured using C_T values (PCR cycle at detection threshold), which are $\log_2$ values. Expression was normalized to that of *PPIA* (endogenous control) and analyzed according to the relative quantification method as $\Delta C_T = C_T$ (*PPIA*) – C_T (target). The fold difference between any two samples was calculated as $2^{(\Delta CT1 - \Delta CT2)}$.

Genotyping. Genotyping was performed using custom-designed TaqMan assays (**Supplementary Fig. 10** and **Supplementary Table 1**) with Genotype Master Mix (Qiagen) on the ABI SDS7700 under standard conditions. Genotyping was performed blinded to clinical phenotypes. Sanger sequencing of HapMap samples and subsets of clinical samples showed complete concordance with the TaqMan genotyping results.

Clinical and epidemiological studies. Studies had been approved by the institutional review boards of the participating institutions, and all participants gave informed consent for genetic testing. Ancestral designation was self-reported in all studies. Full details of the studies are provided in the **Supplementary Note**.

Statistical analysis of genetic association. All statistical comparisons of ss469415590 and rs12979860 in *IFNL4* were limited to subjects successfully genotyped at both variants. The Kruskal-Wallis test was used to compare median HCV RNA levels between genotypes for each variant (TT/TT at ss469415590 versus ΔG/ΔG at ss469415590). The mean HCV RNA levels in each of the three ss469415590 genotype groups (ΔG/ΔG, ΔG/TT and TT/TT) were compared with the levels in the respective rs12979860 genotype groups (TT, CT and CC). The global null hypothesis assumed no difference in decreases of mean HCV RNA values in ss469415590 genotype groups compared to corresponding rs12979860 genotype groups. Statistical significance of these three mean differences was determined using a Wald statistic (3 degrees of freedom) for which the variance was computed using a bootstrap variance estimate based on resampling subjects with replacement (1,000 resamplings). For dichotomous outcomes (SVR or spontaneous clearance) the ORs and accompanying *P* values (Wald χ-square) were calculated using proc logistic (SAS 9.2 TS2M3). The differences in AUC[58] values were evaluated with *P* values on the basis of a χ-square test (1 degree of freedom) that used a bootstrap variance estimate computed by resampling subjects with replacement and then repeating the AUC computations for each bootstrap sample.

56. Wakita, T. *et al.* Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.* **11**, 791–796 (2005).
57. Dickensheets, H.L., Venkataraman, C., Schindler, U. & Donnelly, R.P. Interferons inhibit activation of STAT6 by interleukin 4 in human monocytes by inducing SOCS-1 gene expression. *Proc. Natl. Acad. Sci. USA* **96**, 10800–10805 (1999).
58. Pepe, M.S. *The Statistical Evaluation of Medical Tests for Classification and Prediction* (Oxford University Press, Oxford, 2003).

doi:10.1038/ng.2521