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

Permalink
https://escholarship.org/uc/item/9mj175zw

Journal
Nature genetics, 45(2)

ISSN
1546-1718

Authors
Prokunina-Olsson, Ludmila
Muchmore, Brian
Tang, Wei
et al.

Publication Date
2013-01-06

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

Ludmila Prokunina-Olsson¹, Brian Muchmore¹, Wei Tang¹, Ruth M. Pfeiffer², Heiyoung Park³, Harold Dickensheets⁴, Dianna Hergott¹, Patricia Porter-Gill¹, Adam Mumy¹, Indu Kohaar¹, Sabrina Chen⁵, Nathan Brand¹, McAnthony Tarway¹, Luyang Liu¹, Faruk Sheikh⁴, Jacque Astemborski⁷, Herbert L. Bonkovsky⁸, Brian R. Edlin⁹,¹⁰, Charles D. Howell¹¹, Timothy R. Morgan¹²,¹³, David L. Thomas⁷,¹⁴, Barbara Rehermann³, Raymond P. Donnelly⁴, and Thomas R. O’Brien⁵

¹Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, NCI/NIH, Bethesda, MD 20892, USA ²Biostatistics Branch, Division of Cancer Epidemiology and Genetics, NCI/NIH, Bethesda, MD 20892, USA ³Immunology Section, Liver Diseases Branch, National Institute for Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA ⁴Division of Therapeutic Proteins, Center for Drug Evaluation & Research, Food and Drug Administration, Bethesda, MD 20892, USA ⁵Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI/NIH, Bethesda, Maryland 20892, USA ⁶Information Management Services, Silver Spring, Maryland ⁷Department of Epidemiology, The Johns Hopkins School of Hygiene and Public Health, Baltimore, MD, USA ⁸Department of Medicine, Carolinas Medical Center, Charlotte, NC 28203, USA ⁹National Development and Research Institutes, New York, NY 10010 ¹⁰Center for the Study of Hepatitis C, Weill Medical College of Cornell University, New York, NY, USA ¹¹Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland, USA ¹²Gastroenterology Service, VA Long Beach Healthcare System, Long Beach, CA 90822 ¹³Division of Gastroenterology, University of California - Irvine, Irvine, CA, USA ¹⁴Division of Infectious Diseases, Johns Hopkins University, Baltimore, Maryland 21205, USA

**SUMMARY**

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 dsRNA to...
mimic HCV infection. Upstream of IFNL3 (IL28B) on chromosome 19q13.13, we discovered a novel, transiently induced region that harbors dinucleotide variant ss469415590 (TT/ΔG), which is in high linkage disequilibrium with rs12979860, a genetic marker strongly associated with HCV clearance. ss469415590-ΔG is a frame-shift variant that creates a novel primate-specific gene, designated interferon lambda 4 (IFNL4), which encodes a protein of moderate similarity with IFNL3. Compared to rs12979860, ss469415590 is more strongly associated with HCV clearance in individuals of African ancestry, whereas it provides comparable information in Europeans and Asians. Transient over-expression of IFNL4 in a hepatoma cell line induced STAT1/STAT2 phosphorylation and 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 hepatitis C virus (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 cancer3. Success of CHC treatment with pegylated interferon-α plus 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 is subject to numerous side effects of pegIFN-α/RBV4 and DAA treatment. If this treatment fails, there is an increased risk of selection of resistant HCV strains that may compromise future treatment options5,6.

Recent genome-wide association studies (GWAS) have identified single nucleotide polymorphisms (SNPs) rs12979860 and rs8099917 on chromosome 19q13.13 near IFNL3 (IL28B) gene as markers associated with both spontaneous HCV clearance7,8 and response to pegIFN-α/RBV treatment7–11. Within this region reside the three interferon-λ genes, IFNL1, IFNL2 and IFNL3 (formerly IL29, IL28A and IL28B, respectively), which encode the type-III IFNs12,13. Type-I IFNs (IFN-α and IFN-β) and type-III IFNs (IFN-λ) induce antiviral activity and suppress HCV replication in vitro14,15 and in vivo16 by activation of the JAK/STAT pathway and up-regulation of interferon-stimulated genes (ISGs)14,17,18. The molecular phenotype of this genetic association remains unclear. The GWAS markers have not been consistently associated with hepatic IFNL3 mRNA expression19–21 and a non-synonymous IFNL3 variant, rs8103142 (Arg70Lys), 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 protein22.

A genetic variant creates a novel interferon protein (IFNL4) of similarity with IFNL3

We sequenced mRNA (RNA-seq) from primary human hepatocytes (PHH) treated with 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. The hepatocytes were treated with PolyI:C for 0, 1, 2, 4, 8 or 24 hours, and induction of the IFN-λ genes (IFNL1-3) was confirmed by TaqMan expression analysis prior to RNA-seq. An analysis of RNA-seq data that focused on a 150-Kb region around 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 hours of PolyI:C activation (Fig. 1a).

We also observed transient activation of a novel transcribed region upstream of IFNL3, with the highest levels of expression detected at 2 and 4 hours (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

Nat Genet. Author manuscript; available in PMC 2013 October 09.
start site, followed by a unique protein translation start site 277 bp downstream. Within the first exon (at amino acid 22), we detected a novel compound dinucleotide variant, denoted ss469415590 TT/ΔG, comprised of a one-base insertion/deletion (indel) polymorphism (T/ΔG, rs67272382) and a one-base substitution variant (T/G, rs74597329). Using PolyIC-stimulated PHH from 5 additional liver donors and primer presented 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 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 are alternative splicing forms of IFNL3 or fusions. However, the presence of a CTCF transcriptional insulator site23,24 between the two transcribed regions (Fig. 1b), the results of the RACE experiments and the failure to generate an RT-PCR product between IFNL3 and the novel transcribed region, confirmed their independence. 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 decay25. The remaining six transcripts were predicted to produce full-length proteins of: 143 amino acids (aa; p143) and 124 aa (p124) from transcripts with the ss469415590-TT allele; 179 aa (p179), 170 aa (p170), 131 aa (p131) and 107 aa (p107) from transcripts with the ss469415590-ΔG allele (Fig. 1c). A global protein BLAST search found homology only for p179, with 29.1% aa identity and 40.8% aa similarity with IFNL3. However, the p179 and IFNL3 cDNA sequences were not similar enough to be aligned using a BLAST b2seq tool. Based on the protein sequence homology with type-III IFNs (Table 1), we designated p179 as interferon analog protein (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 interaction of IFNL3 and other type-III IFNs with their primary receptor, IFNLR1 (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 IFNL receptor complex, IL10R226–29.

**Association of genetic variants within 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 the HapMap30 (Supplementary Table 3) and the 1000 Genomes project31 data (Supplementary Table 4) showed that the IFNL4-creating ss469415590-ΔG allele is perfectly correlated with the unfavorable rs12979860-T allele in Asians (CHB+JPT sets, r²=1.00) and well correlated in Europeans (CEU set, HapMap, r²=0.92; 1000 Genomes, r²=0.83). In Africans, however, this correlation is only moderate (YRI set, HapMap, r²=0.71; 1000 Genomes, r²=0.65), even though rs12979860 is the best surrogate for ss469415590 of all markers within the 100-Kb region. Correlation between ss469415590 and rs8099917 is high in Asians (r²=0.91), moderate in Europeans (r²=0.44), but very low in Africans (r²=0.008) (Supplementary Table 3).

We assessed the association of ss469415590 and rs12979860 with HCV clearance in 1,436 African-American and 1,480 European-American individuals from four studies. In VirahepC32 and HALT-C33, we evaluated response to pegIFN-α/RBV therapy in patients 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 racial 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%; African-American, 7%. We evaluated spontaneous
HCV clearance in injection drug users enrolled in two studies, UHS and ALIVE (Supplementary Table 6). The decrease in HCV RNA during the first 28 days of treatment is a powerful early predictor of ultimate treatment response that is strongly associated with rs12979860 genotype. In African-American Virahep-C participants, the decline in HCV RNA levels after 28 days of treatment was more strongly associated with ss469415590 genotype than with rs12979860 genotype (p=0.015, difference in mean values; Fig. 3; 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 patients 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 among 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 rs12979860 (0.58) (Table 3). In the ALIVE study, which has higher percentage of HCV/HIV co-infected patients, the AUC values were similar for rs12979860 (0.64) and ss469415590 (0.64; Table 3).

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, 8, 9, 10). Taken as a whole, our results show that among 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, while these variants are similarly informative in European-Americans.

By sequencing IFNL4 in 270 HapMap samples we annotated three non-synonymous variants, rs73555604 (Cys17Tyr) in exon 1, rs142981501 (Arg60Pro) and rs117648444 (Pro70Ser) in exon 2, as well as four synonymous variants, rs150891559 (Ala11Ala) and rs4803221 (Ser30Ser) in exon 1, and rs12971396 (Ser149Ser) and rs137902769 (Ser175Ser) in exon 5 (Fig. 2, Supplementary Fig. 2, Supplementary Tables 11, 12). Based on a haplotype analysis of 16 markers from the 8-Kb IFNL3/IFNL4 region, we identified eight markers that captured all haplotypes present in HapMap sets (Supplementary Table 13). These eight markers were also tested in European-American and African-American patients from Virahep-C. In all populations, the unique favorable haplotype included the ss469415590-TT allele, which abrogates 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; these two haplotypes include minor alleles of either non-synonymous variants rs73555604 or rs11764844. It is possible, therefore, that these variants modify the risk in carriers of the unfavorable ss469415590-ΔG allele and are the source of haplotype heterogeneity previously reported in Europeans, however, data from Virahep C are too sparse to confirm this finding (Supplementary Table 13).

**IFNL4 induces expression of ISGs**

We evaluated functional properties of the 6 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 6 isoforms or treated with recombinant IFN-α IFNL3 and IFNL4 (the latter two produced in the sfs9 baculoviral expression system). Only transfection with IFNL4 expression construct, as well as treatment with IFN-α or IFNL3, induced activation of an interferon stimulated response element reporter (ISRE-Luc), which contains STAT1/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 IFNL4 expression constructs generating proteins either with a Halo-tag or a FLAG-tag (Supplementary Fig. 3). Similarly, only transient transfection with IFNL4 construct decreased HCV RNA replication in hepatoma cells stably expressing a subgenomic luciferase-expressing hepatitis C virus replicon (Fig. 4d), and induced STAT1 and STAT2 phosphorylation (Fig. 4e). Transfection with IFNL4 activated the ISRE reporter in HepG2 and HEK-293T cells, but not in HeLa cells (Supplementary Fig. 4).

The recombinant IFNL4 protein expression was detectable in cells and cell lysates of transfected HepG2 and 293T cells, by confocal imaging and Western blots with antibodies specific for IFNL4 and tag proteins (Supplementary Fig. 5, 6). We also detected weak expression of recombinant IFNL4 in media of transfected HepG2 cells, but not of 293T cells (Supplementary Fig. 7). In PolyI:C-stimulated PHH from liver donors not infected with HCV endogenous protein expression of IFNL4 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/T-T genotype), we observed low IFNL4 expression even without PolyI:C treatment or in-vitro HCV infection (hour 0 time point). Though 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 IFNL4 or an empty vector (mock) (Supplementary Fig. 8a) and found that the top canonical pathways induced by IFNL4 are related to the activation of interferon signaling (Supplementary Fig. 8b). We validated the RNA-seq results by qRT-PCR analysis and showed that IFNL4 induced expression of many ISGs in a pattern similar to that induced by IFN-α and IFNL3 (Supplementary Fig. 8c). Previously, genes in these pathways have been shown to be expressed at higher levels in pre-treatment liver biopsies from HCV-infected patients who do not respond to pegIFN-α/RBV treatment; these patients tend to carry the unfavorable genotypes of rs12979860 and rs8099917 which mark the ss469415590-ΔG allele that produces IFNL4. To mimic this clinical phenotype, we transfected HepG2 cells with mock 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, IFIH1/MDA5, OAS1, MX1, DHX58/RIG-I) in a pattern similar to that induced by IFN-α and IFNL3 (Supplementary Fig. 8d). Furthermore, treatment by 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 chemokine CCL5 (RANTES) and the proto-oncogene FOS 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.

Conclusions

We have identified a novel inducible human protein-coding gene, IFNL4, which is related to, but distinct from known IFNs and other class-2 cytokines. The 179 aa open-reading frame of IFNL4 transcript is created by a common deletion frame-shift 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 for predicting response to pegIFN-α/RBV treatment of CHC. Within IFNL4 we identified three non-synonymous variants, rs73555604 (Cys17Tyr), rs142981501 (Arg60Pro) and rs11764844 (Pro70Ser), present on haplotypes with the ss469415590-ΔG allele. The role of these variants on IFNL4 biological function and their impact on HCV clearance in different populations should be further explored.

Analysis of genomic sequences of 45 species available for IFNL4 region in the UCSC genome browser showed that the unfavorable, IFNL4-generating ss469415590-ΔG allele is an ancestral variant present in all the species. The existence of IFNL4 protein could be predicted only in the genomes of macaques (marmoset and rhesus), orangutan, chimpanzee and humans (Supplementary Fig. 9). The beneficial insertion ss469415590-TT allele appears to be a recently derived variant, which became common in all human populations (93% in Asians, 68% in Europeans and 23% in Africans, based on HapMap samples), suggesting positive selection for this allele. Introduction of frame-shifts is considered to be an evolutionary mechanism for the rapid emergence of new proteins49,50, 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 aa 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, is currently under investigation. IFNL4 and IFNL3 share similarity 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 binding of IFNLR1. We also found that the IFNL4-caused pre-activation 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 PHH, where it was induced by PolyI:C, IFN-α (but not with IFNL3) 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, HEK293T, 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 patients with chronic hepatitis C who carry rs12979860-T, which marks the ss469415590-ΔG allele, have somewhat higher hepatic expression of ISGs before treatment, but poorer ISG response to pegIFN-α/RBV treatment19,41–44. The rs12979860-T variant has also been associated with lower HCV RNA levels in the absence of treatment9,51. 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. Together, these data suggest that IFNL4 induces weak expression of ISGs that provides some antiviral effect response which might lower the HCV load, but 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 treatment52. This genotype has also been associated with response to IFN-α-based treatment of chronic hepatitis B virus (HBV) infection in some studies53,54. Furthermore, IFN-α therapy is used for a number of other clinical conditions, including some forms of cancer55. Thus, therapeutic inhibition of IFNL4 might represent a novel biological target for the treatment of HCV and HBV infection and possibly other diseases, and IFNL4 genotype could be used to select patients for this therapy.
ONLINE METHODS

Cells
Fresh primary human hepatocytes (PHH) from HCV un-infected liver donors were purchased from Lonza or Celsis. PHH received in suspension were cultured overnight in collagen-coated plates or chamber slides (BD Biosciences), unattached PHH were removed before treatment. Hepatoma HepG2 cells, cervical carcinoma HeLa and embryonic kidney 293T cells (all from ATCC) were cultured in DMEM with 10% FBS. Hepatoma Huh7-Lunet cells harboring a subgenomic, luciferase-expressing HCV JFH1 replicon were cultured as previously described. The PHH were treated with 50 μg/ml of PolyI:C (Imgenex), 100 ng/ml of IFN-α (PBL Interferon Source) or IFNL3 (custom); cell lines were treated with 10 ng/ml of IFN-α, IFNL3 or IFNL4 (custom).

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

HCV infection of PHH
PHH (0.35x10^5/well) were infected with HCV (JFH1 strain) for 0, 6 or 24 hours in collagen-coated 24-well plates or chamber slides at a multiplicity of infection (M.O.I.) of 2.

Antiviral assays
Huh7-Lunet cells were transfected with IFNL4, p131 or p107 constructs or an empty GFP vector (mock) in a 48-well plate and luciferase expression was measured 48 hours post-transfection.

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

Western blotting
Proteins were resolved on 4–12% tris-glycine SDS-PAGE gels (Life Technologies). Detection was done using the custom anti-IFNL4 mouse and rabbit monoclonal antibodies, rabbit anti-Halo antibody (Promega), and secondary goat anti-rabbit (sc-2030) or goat anti-mouse (sc-2031) antibodies with IgG-HRP (Santa Cruz). The signals were detected with ECL Plus Western blotting detection system (GE Healthcare Life Sciences).

Analysis of STAT1/STAT2 phosphorylation
HepG2 cells in 6-well plates were untreated or transfected with expression constructs or empty Halo-tag vector, or were treated for 1 hour at 37°C with 50 ng/ml of recombinant IFNL3. Equal amounts (50 μg/lane) of whole-cell lysates prepared 48 hours post-transfection were used for analysis by Western blotting. Detection was performed as previously described with rabbit anti-phospho-Tyr701-STAT1 (Cell Signaling Technology) and rabbit anti-phospho-Tyr689-STAT2 (Millipore) antibodies. The blots were stripped and re-probed with rabbit anti-STAT1 and anti-STAT2 antibodies (Santa Cruz Biotechnology) to measure the levels of total STAT1 and STAT2 proteins.
RNA-sequencing

Total RNA (1 ug) from PHH or HepG2 cells was used for library preparation with TruSeq PolyA kit (Illumina). Sequencing with Genome Analyzer (GAII) generated 47 million of 107 bp paired-end sequencing reads per sample. The TopHat v1.2.0 settings were changed to allow multiple read alignments (up to 10 regions, UCSC hg19) and 3 nucleotide mismatches per each 25-bp segment. Results were viewed with the UCSC genome browser http://genome.ucsc.edu/ and the Integrative Genomics Viewer (IGV) software: broadinstitute.org/igv/.

Identification and cloning of novel splicing forms

Rapid amplification of cDNA ends (5′ and 3′ RACE) and cloning of full-length open reading frames were performed with SMARTer RACE cDNA kit (Clontech,) using a pool of DNase-I treated RNA samples from PolyI:C-treated PHHs from several liver donors (Supplementary Table 1 for primers). PCR reactions were performed with AmpliTaq Gold 360 Master Mix (Life Technologies) and 360 GC Enhancer (Life Technologies) using the touchdown PCR program: 10 minutes at 95 °C, 20 cycles of touchdown (30 seconds at 95 °C, 45 seconds from 70 to 60°C decreasing by 1 °C after each 2 cycles, 45 seconds at 72 °C); 25 additional cycles (30 seconds at 95 °C, 45 seconds at 60 °C and 45 seconds at 72 °C); and final extension time of 7 minutes at 72 °C. Gel-purified PCR fragments were cloned into a C-terminal pFC14A–Halo tag expression vector (Promega) and sequenced for validation. IFNL3-Halo expression construct was generated using the same approach. p179 was also cloned into a pcDNA3.1-FLAG-tagged expression vector.

Production of recombinant proteins

IFNL4 and IFNL3 bacmids were generated in pFastBac C-terminal His-tag vector (Life Technologies) and expressed in a sf9 baculoviral strain. Using the anti-His-tag antibody (Sigma), expression of IFNL3 was detected in cell media, which was used for protein purification. Expression of p179 was not detectable in cell media and protein was purified from the cell pellet after cultivation of cells for 3–5 days in 2 liters of SF-900 III medium (Life Technologies). Proteins were first purified on HisTrap 5-ml nickel column followed by size-exclusion chromatography preparative column TSK G3000pw of 21.5×600 mm (Tosoh). The purified protein fractions were concentrated and dialyzed into a buffer. High protein purity (>90%) was confirmed by Coomassie staining and Western blot analyses with anti-His antibody (#H1029, Sigma, 1:3000 dilution), anti-IFNL3 and custom mouse and rabbit monoclonal anti-p179 antibodies. The IFNL3 and p179 proteins were custom-produced by Protein One.

Development of anti-IFNL4 antibodies

A mouse monoclonal antibody was custom-developed for a synthetic peptide KALDRYEEEALSWGQRNCSFRPRDMSPS corresponding to amino acids 44–74 of p179 protein, by Precision Antibody. A rabbit monoclonal antibody was custom-developed for a synthetic peptide PGSSRKVPGAQKRRHKPRRDSPRC corresponding to amino acids 128–152 of p179 protein, by Epitomics.

Evaluation of biological activity of novel proteins

Luciferase Cignal 45-Pathway Finder Reporter Arrays were used in HepG2 cells according to instructions (Qiagen, full list of reporters tested is available at http://www.sabiosciences.com/reporter_assay_product/HTML/CDA-901L.html). Cells were transfected with expression constructs for p179, p170, p143, p131, p124 and p107 or treated for 24 h with purified recombinant proteins ~10 ng/ml of IFN-α or IFNL3 and/or IFNL4-p179. Validation was performed with an individual interferon-stimulated response element.
Global analysis of transcriptome and pathway analysis

HepG2 cells were mock-transfected with an empty Halo-tag vector or with IFNL4-Halo expression construct. High-quality RNA (RIN ~10) prepared from transfected cells (48 hours) was used for sequencing with HiSeq 2000 (Illumina), generating ~300M reads per sample. Standard analysis identified 535 transcripts with >2-fold difference in expression and an FDR<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 mock, IFNL4, p107, p131 constructs or/and treated with 10 ng/ml of IFN-α or IFNL3, in 4 biological replicates. mRNA expression in all samples was evaluated with pathway-based RT² Profiler PCR arrays, according to instructions (Qiagen).

Confocal Imaging

PHH cultured on collagen-coated slides were treated with 50 μg/ml PolyI:C for 0 h (untreated), 2, 4, 8 and 24 hrs. HepG2 cells were transiently transfected with IFNL4-Halo expression construct. Cells were fixed for 20 min with 4% formaldehyde (Sigma) in PBS, permeabilized with 0.5% TritonX 100 (Sigma) for 5 min, blocked with 4% BSA (Sigma) and then incubated overnight at 4°C with primary antibodies: mouse monoclonal anti-IFNL4 ab (custom), rabbit α-tubulin ab (ab-15246, Abcam), rabbit anti-Halo-tag ab (Promega), mouse α-tubulin ab (ab-7291, Abcam, all at 1:1000 dilution. Secondary antibodies were donkey anti-rabbit or anti-mouse Alexa Fluor 594 and donkey anti-rabbit and anti-mouse Alexa Fluor 488, all at 1:1000 dilutions (Life Technologies). Slides were covered with mounting media (Prolong Gold Antifade Reagent with DAPI). Immunofluorescent images were obtained with a confocal laser-scanning microscope (LSM 510 META, Carl Zeiss).

Analysis of IFNL4 mRNA expression

All expression assays (Supplementary Table 1 and Supplementary Fig. 10) were purchased from Life Technologies. Expression analysis was performed with gene expression master mix (Life Technologies) on DNAseI-treated RNA samples on ABI SDS 7700 instrument. Expression was measured in Ct values (PCR cycle at detection threshold), which are log 2 values. Expression was normalized to PPIA (endogenous control) and analyzed according to relative quantification method, as ΔCt=Ct_PPIA−Ct_target. Fold difference between any two samples can be calculated as fold=2^((ΔCt_1−ΔCt_2).

Genotyping

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

Clinical and epidemiological studies

The studies had been approved by the institutional review boards of the participating institutions and all patients gave informed consent for genetic testing. Ancestral designation
was self-reported in all studies. Full details of the studies are provided in a Supplementary Note included in Supplementary Materials.

**Statistical analysis of genetic association**

All statistical comparisons of *IFNL4* ss469415590 and rs12979860 were limited to subjects successfully genotyped for both variants. The Kruskal-Wallis test was used to compare median HCV RNA levels between genotypes for each variant (e.g., ss469415590-TT/TT versus ss469415590-ΔG/ΔG). The mean HCV RNA levels in each of the three ss469415590 genotype groups (i.e., ΔG/ΔG, ΔG/TG, TT/TT) were compared with the respective rs12979860 genotype groups (TT, CT, CC). The global null hypothesis assumed no difference in decreases of mean HCV RNA values in ss469415590 compared to corresponding rs12979860 genotype groups. Statistical significance of these three mean differences was based on a Wald statistic (3 df) for which the variance was computed using a bootstrap based on re-sampling subjects with replacement (1000 re-samplings). For dichotomous outcomes (e.g., SVR, spontaneous clearance) the odds ratios and accompanying p-values (Wald chi-square) were calculated using proc logistic (SAS 9.2 TS2M3). The differences in the area under the receiver operating characteristic curve (AUC) values were evaluated with p-values based on a chi-square test (1 df) that used a bootstrap variance estimate computed by re-sampling subjects with replacement and then repeating the AUC computations for each bootstrap sample.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

This research was supported by the Intramural Research Program of the 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, NIH grants R01-DA12109, R01-DA13245 and R01-DA16159 (B.E.); NCI contracts N02-CP-91027 and N01-CO-12400 (B.E.), Substance Abuse and Mental Health Services Administration grant H79-T12103 (B.E.), and the NCI Director’s Innovation Award (L.P.-O.).

The Virahep-C and HALT-C studies were conducted by the Virahep-C and HALT-C Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The data and samples from the Virahep-C and HALT-C studies reported here were supplied by the NIDDK Central Repositories. This manuscript was not prepared in collaboration with the Virahep-C and HALT-C study groups and does not necessarily reflect the opinions or views of the Virahep-C and HALT-C studies, the NIDDK Central Repositories, or the NIDDK.

We thank NCI Sequencing Core Facility for help with RNA-seq. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

**References**


54. Lampertico P, et al. IL28B polymorphisms predict interferon-related HBsAg seroclearance in genotype D HBeAg-negative patients with chronic hepatitis B. Hepatology. 2012


Figure 1. Identification of a novel transcribed region upstream of IFNL3 gene
RNA-seq in primary human hepatocytes (PHH) treated with 50 ug/ml PolyI:C for 0, 1, 2, 4, 8 or 24 hours. RNA-seq plot of the 150-Kb region in USCS browser shows expression of IFNL1, IFNL2 and IFNL3, and a novel transcribed region upstream of IFNL3. The number of reads (depth) corresponds to the level of mRNA expression.

a. Detailed view of the RNA-seq results for IFNL3 and the novel transcribed region. A CTCF transcriptional insulator (ENCODE data^24) between these regions indicates their independence.

b. Splicing architecture of the ten novel transcripts (NCBI accession numbers are presented in Supplementary Table 2). The GWAS marker, rs12979860, is located within the first intron, while a novel marker, ss469415590-TT/ΔG, is located within the first exon, common for all transcripts. Transcription and translation start sites are marked by black and blue arrows; open reading frames are shaded in blue. * -transcripts which carry premature stop codons and are likely to be eliminated by nonsense-mediated decay. Arrows indicate location of primers (Supplementary Table 1) used to generate PCR-products presented on panel d.

c. PCR results in PHH cDNA using primers from panel c. No distinct PCR product is expected in the TT/TT sample; in the ΔG/ΔG sample these primers capture transcripts producing proteins of 179, 131 and 107 aa, but not of 170 aa; a transcript producing a 93 aa protein fragment is expected to be degraded. IFNL3, IFNL1 and PPIA (endogenous control) were measured in the same samples. Similar amounts of DNaseI-treated high quality RNA was used for all the reactions.
Figure 2. Protein sequence analysis
ClustalW protein sequence alignment for IFNL4 (p179), p131, p107, IFN-β (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 mature IFNL3 protein (without leader peptide). Indicated: cysteins conserved between IFNL4 and all other IFNL proteins (C16, C50, C115, C148 and C174), interaction sites of IFNLs with their first common receptor IFNLR1 (aa 27, 33, 34, 36, 37, 44, 53, 155, 158); and interaction sites with their second receptor, IL10R2 (aa 97, 100).
Figure 3. Median decrease in HCV RNA (log$_{10}$ IU/ml) in African-American participants in Virahep-C study during the first 28 days of treatment with pegIFN-α/RBV. P=0.015 for comparison of mean differences in HCV RNA levels at day 28 for each of the three genotype groups for ss469415590 with the respective group for rs12979860.
Figure 4. Analysis of biological activity of 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 (mock), or treated with 10 ng/ml of recombinant purified IFN-α, IFNL3, IFNL4, or with PBS (mock). All results represent mean values of two independent biological transfection/treatment replicates, with standard deviations.

b. Results of transient transfection of the IFNL4, p131 and p107 expression constructs and treatment with recombinant purified IFN-α and IFNL3 proteins in HepG2 cell line transiently co-transfected with ISRE-Luc reporter. The results are normalized to transfection with an empty vector (mock) and represent mean values of 8 biological replicates, with standard errors.

c. Results of transient transfection of the IFNL4, p131 and p107 expression constructs in HepG2 cell line stably expressing the ISRE-Luc reporter. The results are normalized to transfection with an empty vector (mock) and represent mean values of 11 biological replicates, with standard errors.

d. Test for antiviral effects of the IFNL4, p131 and p107 expression constructs transiently transfected into Huh7-Lunet cells stably expressing a subgenomic, luciferase-expressing

Nat Genet. Author manuscript; available in PMC 2013 October 09.
HCV replicon (HCV-Luc), compared to an empty vector (mock). Results represent mean values of 4 biological replicates, with standard errors.
e. Western blot analysis of Y701-STAT1 and Y689-STAT2 phosphorylation in HepG2 cells transiently transfected with expression constructs for the six protein isoforms. All Halo-tag expression constructs produce proteins detectable with an antibody for the Halo-tag; a rabbit monoclonal anti-IFNL4 antibody recognizes IFNL4 (p179) as well as its nonfunctional isoforms p131 and p107.
Figure 5. Confocal imaging of IFNL4 in primary human hepatocytes (PHH) from liver donors with different ss469415590 genotypes

Confocal imaging in PHH treated with 50 μg/ml of PolyI:C for 0, 2, 4, 8 or 24 hours. Red staining – IFNL4, green – cytoskeleton (α-tubulin), blue – nuclei. Confocal imaging of endogenous IFNL4 expression in the same ss469415590-ΔG/TT sample after in-vitro infection with JFH1-HCV strain is presented in Supplementary Fig. 6.
**Table 1**

Comparisons between p179 (IFNL4) and selected members of class-2 cytokine family

<table>
<thead>
<tr>
<th>Protein</th>
<th>p179 (IFNL4)</th>
<th>IFNL3</th>
<th>IFNL1</th>
<th>IFNL2</th>
<th>IFN-α</th>
<th>IL10</th>
<th>IL22</th>
<th>IFN-β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>IFN-III</td>
<td>IFN-III</td>
<td>IFN-III</td>
<td>IFN-III</td>
<td>IFN-III</td>
<td>IFN-I</td>
<td>IFN-10-family</td>
<td>IFN-10-family</td>
</tr>
<tr>
<td>mRNA</td>
<td>JN806234</td>
<td>NM_172139</td>
<td>NM_172140</td>
<td>NM_172138</td>
<td>NM_002175</td>
<td>NM_000572</td>
<td>NM_020525</td>
<td>NM_002176</td>
</tr>
<tr>
<td>Protein ID</td>
<td>AFQ38559</td>
<td>A2BDE1</td>
<td>Q8IU54</td>
<td>Q45KQ8</td>
<td>P05014</td>
<td>Q6FGW4</td>
<td>Q9GZ6</td>
<td>Q5VWC9</td>
</tr>
<tr>
<td>Size, aa</td>
<td>179</td>
<td>196</td>
<td>200</td>
<td>200</td>
<td>189</td>
<td>178</td>
<td>179</td>
<td>187</td>
</tr>
<tr>
<td>Exons</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Identity with p179, aa, n (%)</td>
<td>--</td>
<td>52 (29.1)</td>
<td>48 (26.8)</td>
<td>47 (26.3)</td>
<td>26 (14.5)</td>
<td>24 (13.4)</td>
<td>20 (11.1)</td>
<td>20 (11.1)</td>
</tr>
<tr>
<td>Similarity with p179, aa, n (%)</td>
<td>--</td>
<td>73 (40.8)</td>
<td>66 (36.8)</td>
<td>70 (39.1)</td>
<td>47 (26.2)</td>
<td>40 (22.3)</td>
<td>45 (25.1)</td>
<td>44 (24.5)</td>
</tr>
</tbody>
</table>
Comparison of rs12979860 and ss469415590 genotypes among African-American participants of Virahep-C study and HALT-C Trial for associations with response to treatment with pegIFN-α/RBV

<table>
<thead>
<tr>
<th>Variant</th>
<th>Genotype</th>
<th>rs12979860</th>
<th>ss469415590</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(n = 169)</td>
<td>(n = 169)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Week 20/24</td>
<td>End-of-Treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Response Rate (%)</td>
<td>OR</td>
</tr>
<tr>
<td>TT</td>
<td>57</td>
<td>40.4</td>
<td>Ref.</td>
</tr>
<tr>
<td>CT</td>
<td>93</td>
<td>50.5</td>
<td>1.51</td>
</tr>
<tr>
<td>CC</td>
<td>19</td>
<td>63.2</td>
<td>2.53</td>
</tr>
<tr>
<td>ΔG/ΔG</td>
<td>71</td>
<td>35.2</td>
<td>Ref.</td>
</tr>
<tr>
<td>ΔG/TT</td>
<td>82</td>
<td>56.1</td>
<td>2.35</td>
</tr>
<tr>
<td>TT/TT</td>
<td>16</td>
<td>68.8</td>
<td>4.05</td>
</tr>
<tr>
<td>TT</td>
<td>61</td>
<td>9.8</td>
<td>Ref.</td>
</tr>
<tr>
<td>CT</td>
<td>72</td>
<td>16.7</td>
<td>1.83</td>
</tr>
<tr>
<td>CC</td>
<td>11</td>
<td>45.5</td>
<td>7.64</td>
</tr>
<tr>
<td>ΔG/ΔG</td>
<td>68</td>
<td>8.8</td>
<td>Ref.</td>
</tr>
<tr>
<td>ΔG/TT</td>
<td>68</td>
<td>19.1</td>
<td>2.44</td>
</tr>
<tr>
<td>TT/TT</td>
<td>8</td>
<td>50.0</td>
<td>10.34</td>
</tr>
</tbody>
</table>

In Virahep-C, response is at 24 weeks after initiation of pegIFN-α/RBV treatment; in HALT-C, response is at 20 weeks after initiation of pegIFN-α/RBV treatment; SVR - sustained virological response; OR - odds ratio.
Table 3

Comparison of rs12979860 and ss469415590 genotypes among African-American participants of the Urban Health Study (UHS) and ALIVE for associations with spontaneous clearance of hepatitis C virus (HCV) infection

<table>
<thead>
<tr>
<th>Variant</th>
<th>Genotype</th>
<th>Urban Health Study (n=459)</th>
<th>ALIVE (N=664)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Genotype</td>
<td></td>
</tr>
<tr>
<td>rs12979860</td>
<td>TT</td>
<td>119</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34.0</td>
<td>38.9</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>183</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.3</td>
<td>46.9</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>48</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.7</td>
<td>14.2</td>
</tr>
<tr>
<td>ss469415590</td>
<td>ΔG/ΔG</td>
<td>137</td>
<td>255</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39.1</td>
<td>43.5</td>
</tr>
<tr>
<td></td>
<td>ΔG/TT</td>
<td>168</td>
<td>265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48</td>
<td>45.2</td>
</tr>
<tr>
<td></td>
<td>TT/TT</td>
<td>45</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12.9</td>
<td>11.3</td>
</tr>
</tbody>
</table>

*Percentage of subjects with indicated genotype; OR - odds ratio; AUC - area under the receiver operating characteristic curve; Urban Health Study \(^{34}\) and ALIVE \(^{35}\) are epidemiological studies of injection drug users who were evaluated for chronic HCV infection on the basis of testing for anti-HCV and HCV RNA.