Supplementary Materials

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

Ludmila Prokunina-Olsson, Brian Muchmore, Wei Tang, Ruth M. Pfeiffer, Heiyoung Park, Harold Dickensheets, Dianna Hergott, Patricia Porter-Gill, Adam Mumy, Indu Kohaar, Sabrina Chen, Nathan Brand, McAnthony Tarway, Luyang Liu, Faruk Sheikh, Jacquie Astemborski, Herbert L. Bonkovsky, Brian R. Edlin, Charles D. Howell, Timothy R. Morgan, David L. Thomas, Barbara Rehermann, Raymond P. Donnelly, Thomas R. O'Brien.

Supplementary Tables

Nature Genetics: doi:10.1038/ng.2521

Supplementary Table 12. Analysis of pair-wise linkage disequilibrium (r^2) within *IFNL3/IFNL4* region in HapMap samples

Supplementary Table 13. Haplotype analysis in the *IFNL3/IFNL4* region in HapMap samples and participants of Virahep-C study

Supplementary Figures

Supplementary Fig. 1. DNA and protein sequence analysis of the regions upstream of the *IFNL2* and *IFNL3* genes

Supplementary Fig. 2. Sequence of *IFNL4* gene and genetic variants found to be polymorphic at least in one population by sequencing of 270 HapMap samples (CEU, YRI and CHB/JPT)

Supplementary Fig. 3. Activation of the Interferon Stimulated Response Element (ISRE) - Luc reporter in HepG2 cells transiently co-transfected with IFNL4 expression constructs carrying either the Halo-tag or a FLAG-tag

Supplementary Fig. 4. Activation of the interferon-stimulated response element (ISRE)-Luc reporter and mRNA expression of receptors for type–I IFNs (*IFNAR1* and *IFNAR2*) and type–III IFNs, *IFNLR1* (*IL28R1)* and *IL10R2*, in human HepG2, 293T and HeLa cell lines

Supplementary Fig. 5. Outline of expression constructs for the 6 protein isoforms

Supplementary Fig. 6. Analysis of IFNL4 expression in HepG2 cells and primary human hepatocytes

Supplementary Fig. 7. Analysis of IFNL4 expression in 293T and HepG2 cells transiently transfected with IFNL4-Halo construct or an empty Halo-tag vector

Supplementary Fig. 8. Functional effects of IFNL4 over-expression in HepG2 cells

Supplementary Fig. 9. Prediction of IFNL4 protein based on genomic sequences of 45 species available in UCSC genome browser

Supplementary Fig. 10. Outline of mRNA transcripts within *IFNL4* region and their detection by RNA-sequencing and an allele-specific expression assay

Supplementary Fig. 11. Quantitative reverse-transcriptase PCR (qRT-PCR) analysis of allele-specific expression of transcripts with ss469415590-TT and ΔG alleles and expression of *IFNL3 (IL28B)*, *IFNL1* (*IL29)* and *PPIA* (endogenous control) in various samples

Supplementary Note. Clinical Studies

Supplementary Table 1. PCR primers and assays

primer	sequence	comments
IFNL4_cloning_F	ATGCGGCCGAGTGTCTGGGCC	additional overhangs were added for cloning into
IFNL4_cloning_R	GAGGCAAGGCCCAGAGTGTGCAG	specific vectors
IFNL4_F_seq_intr1_3'UTR	GTAAGTCACCGCCCAGCCCCTGTGCC	1397-bp amplicon for genotyping by sequencing of IFNL4 variants within intron
IFNL4_R_seq_intr1_3'UTR	CCCATTGACTGAGAGCCTCGCCCGG	1-exon 5
IFNL4_ex1_seq_F	CGAACCAGGGTTGAATTGC	628-bp amplicon for genotyping by sequencing of IFNL4 variants within exon 1
IFNL4_ex1_seq_R	GCACTGCAGACAGGAGTGAG	and 5'UTR
TaqMan expression assays with MGB probes		
ss469415590_IFNL4_F	GCCTGCTGCAGAAGCAGAGAT	ABI expression buffer
ss469415590_IFNL4_R	GCTCCAGCGAGCGGTAGTG	RNA must be DNAseI-treated
ss469415590_IFNL4_VIC (TT allele)	ATCGCAG**AA**GGCC	
ss469415590_IFNL4_FAM (ΔG, allele)	ATCGCAG**C**GGCCC	
IFNL3_F	CGGAAGAGGTTGAAGGTGAC	ABI expression buffer
IFNL3_R	CTCCACCATTGGCTGCAC	RNA must be DNAseI-treated
IFNL3_probe (FAM)	GCCCCAAAAAAGGA	182 bp on DNA, 89 bp on cDNA template
IFNL1	assay Hs00601677_g1, Applied Biosystems	
PPIA (endogenous control)	assay 4326316E, Applied Biosystems	
TaqMan genotyping assays with MGB probes		
ss469415590_IFNL4_F	GCCTGCTGCAGAAGCAGAGAT	standard, Qiagen genotyping buffer
ss469415590_IFNL4_R	GCTCCAGCGAGCGGTAGTG	
ss469415590_IFNL4_VIC (TT, non-risk allele)	ATCGCAG**AA**GGCC	
ss469415590_IFNL4_FAM (dG, risk allele)	ATCGCAG**C**GGCCC	
rs12979860_F	GCCTGTCGTGTACTGAACCA	
rs12979860_R	GCGCGGAGTGCAATTCAAC	
rs12979860_FAM (T, risk allele)	CTGGTTC**A**CGCCTTC	
rs12979860_VIC (C, non-risk allele)	TGGTTC**G**CGCCTTC	
rs8099917	pre-developed assay C__11710096_10, Applied Biosystems	

Supplementary Table 2. **Description of novel transcripts and proteins identified upstream of *IFNL3* gene**

Transcript ID, NCBI accession number	Exons	mRNA	Protein	ss469415590 allele	Annotation
IFNL4, p179, JN806234	5	1636 bp	179 aa	ΔG	*Similarity to IFNL3, full-length protein
p131, JN806225	4	1492 bp	131 aa	ΔG	Similarity to IFNL3, full-length protein
p107, JN806226	3	1420 bp	107 aa	ΔG	Similarity to IFNL3, full-length protein
JN806232	3	1131 bp	93 aa	ΔG	No similarity, protein fragment
p170, JN806233	3	915 bp	170 aa	ΔG	No similarity, full-length protein
JN806227	3	1637 bp	123 aa	TT	No similarity, protein fragment
JN806228	2	1493 bp	75 aa	TT	No similarity, protein fragment
JN806229	1	1421 bp	51 aa	TT	No similarity, protein fragment
p124, JN806230	3	916 bp	124 aa	TT	No similarity, full-length protein
p143, JN806231	3	1132 bp	143 aa	TT	No similarity, full-length protein

*Protein similarity is defined based on global protein BLAST search based on NCBI databases. IFNL4 (p179) has highest homology with IFNL3 – 29.1% amino acid identity and 40.8% amino acid similarity. Protein fragments are open reading frames subjected to nonsense-mediated decay due to the presence of premature stop codons.

Nature Genetics: doi:10.1038/ng.2521

Supplementary Table 3. Allele frequencies of ss469415590, rs12979860 and rs8099917 in HapMap populations

rs8099917 genotypes were downloaded from HapMap (www.hapmap.org), while ss469415590 and rs12979860 were genotyped in all HapMap samples. Pair-wise linkage disequilibrium estimates (r^2) are between ss469415590 and other markers.

Variant, allele	**HapMap populations**					
	West-Africans (YRI) 30 trios		Europeans (CEU) 30 trios		Asians (Chinese/CHB and Japanese/JPT) 90 individuals	
	Frequency %	r^2	Frequency %	r^2	Frequency %	r^2
ss469415590, TT	23.3	-	67.6	-	93.3	-
rs12979860, C	30.0	0.71	67.3	0.92	93.3	1.00
rs8099917, T	97.5	0.008	81.7	0.44	93.9	0.91

Nature Genetics: doi:10.1038/ng.2521

Supplementary Table 4. Analysis of linkage disequilibrium in 1000 Genomes Project.

SNPs from a 100-Kb genomic region in high linkage disequilibrium with ss469415590 (r^2>0.6 in YRI or r^2>0.75 in other populations) in the 1000 Genomes Project reference panel (http://www.1000genomes.org), October 2010 release. In the 1000 Genomes reference panel ss469415590 (T**T**/Δ**G**) is represented by rs74597329 (**T**/**G**). The GWAS marker rs12979860 is shown in bold.

Population	Marker	Distance (bp)	r^2	Allele 1	Allele 1 frequency, %	N, chromosomes
CEU	rs74597329 (ss469415590)	0	1	T (TT)	74.2	120
	rs4803222	198	0.914	G	74.2	120
	rs8113007	3948	0.874	A	73.3	120
	rs688187	6403	0.846	G	70.8	120
	rs4803217	4935	0.837	C	72.5	120
	rs12979860	**368**	**0.832**	**C**	**74.2**	**120**
	rs581930	6032	0.802	C	71.7	120
YRI	rs74597329 (ss469415590)	0	1	T (TT)	28.0	120
	rs12979860	**368**	**0.658**	**C**	**30.5**	**118**
	rs73930703	1642	0.642	C	33.1	118
CHB/JPY	rs74597329 (ss469415590)	0	1	T (TT)	94.2	120
	rs12979860	**368**	**1**	**C**	**94.2**	**120**
	rs12980275	7372	0.85	A	95.0	120
	rs4803221	26	0.764	C	92.5	120
	rs4803222	198	0.764	G	92.5	120

Supplementary Table 5. Characteristics of the Virahep-C and HALT-C participants

Characteristic	Virahep-C				HALT-C			
	African-American N = 169		European-American N = 182		African-American N = 144		European-American N = 741	
Age (Median, IQR)	49.1	45.6 - 52.8	48.1	42.9 - 52.4	51	46.5 - 56.0	49	45.0 - 53.0
Male (N, %)	110	65.1	120	65.9	82	56.9	559	75.4
Ishak Fibrosis Score								
Missing (N, %)	1	0.6					1	0.1
0 (N, %)	18	10.7	19	10.4				
1 (N, %)	44	26	42	23.1				
2 (N, %)	49	29	49	26.9	14	9.7	54	7.3
3-4 (N, %)	49	29	55	30.2	82	56.9	413	55.7
5-6 (N, %)	8	4.7	17	9.3	48	33.3	273	36.8
HCV Genotype 1 (N, %)	169	100	182	100	138	95.8	656	88.5
HCV RNA level (log 10 IU) (median, IQR)	6.4	5.6 - 6.7	6.5	5.7 - 6.8	6.4	6.1 - 6.8	6.5	6.1 - 6.8
Prior Treatment (N, %)								
None	169	100	182	100				
pegIFN-α					32	22.2	211	28.5
pegIFN-α/RBV					112	77.8	530	71.5

IQR - interquartile range

Nature Genetics: doi:10.1038/ng.2521

Supplementary Table 6. Characteristics of the UHS and ALIVE participants

Characteristic	Urban Health Study				ALIVE	
	African-American		European-American		African-American	
	Chronic	Cleared	Chronic	Cleared	Chronic	Cleared
N	350	109	395	162	586	78
Age, median (IQR)	46.0 (42 – 50)	46.0 (42 – 49)	42.0 (36 – 48)	39.5 (33 – 47)	40 (36 – 45)	40 (37 – 43)
Years injection drug use, median (IQR)	27.0 (21 – 32)	25.0 (19 – 31)	23.0 (16 – 29)	20.0 (11 – 29)	14 (7 – 19)	13 (8 – 19)
Male (%)	62.6	60.6	72.4	67.3	76	65
HIV-1-infected (%)	13.1	10.1	12.2	4.9	54	33
Chronic HBV (%)	2.9	9.2	3.3	4.3	3	7

IQR - interquartile range

Supplementary Table 7. Comparison of rs12979860 and ss469415590 genotypes for predicting decrease in HCV RNA load. The median decrease in HCV RNA load (log_{10} IU/ml) is evaluated after 28 days of treatment with pegIFN-α/RBV in European-American (EA) and African-American (AA) participants in Virahep-C. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Variant	Genotype	N	%	HCV RNA Decrease (Median, log_{10} IU/ml)	HCV RNA Difference*	*P* value**
			African-American			
rs12979860	TT	53	32.7	1.05	0.0	
	CT	91	56.2	1.31	0.26	0.031
	CC	18	11.1	2.70	1.65	$1.4x10^{-4}$
ss469415590	ΔG/ΔG	67	41.4	0.87	0.0	
	ΔG/TT	80	49.4	1.54	0.67	$8.4x10^{-5}$
	TT/TT	15	9.3	2.84	1.97	$6.6x10^{-7}$
			European-American			
rs12979860	TT	19	10.8	1.14	0.0	
	CT	78	44.3	1.57	0.43	0.027
	CC	79	44.9	2.99	1.85	$3.4x10^{-7}$
ss469415590	ΔG/ΔG	19	10.8	1.14	0.0	
	ΔG/TT	78	44.3	1.57	0.43	0.027
	TT/TT	79	44.9	2.99	1.85	$3.4x10^{-7}$

* Difference in median HCV RNA level (log 10 IU/ml) compared to referent genotype
** Kruskal-Wallis test

Supplementary Table 8. Comparison of rs12979860 and ss469415590 genotypes for predicting response to pegIFN-α/RBV treatment in European-Americans from Virahep-C study. Response is evaluated as 24-week, end-of-treatment and sustained virological response (SVR) to pegIFN-α/RBV treatment in European-American (n=182) participants in Virahep-C. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Variant	Genotype	N	24-week Response			End-of-Treatment Response			SVR		
			%	OR	p-value	%	OR	p-value	%	OR	*P* value
rs12979860	TT	21	52.4	Ref.		33.3	Ref.		28.6	Ref.	
	CT	78	74.4	2.64	0.056	67.9	4.24	5.7×10^{-3}	44.9	2.03	0.18
	CC	83	80.7	3.81	9.9×10^{-3}	72.3	5.22	1.6×10^{-3}	63.9	4.42	5.4×10^{-3}
ss469415590	ΔG/ΔG	21	52.4	Ref.		33.3	Ref.		28.6	Ref.	
	ΔG/TT	78	74.4	2.64	0.056	67.9	4.24	5.7×10^{-3}	44.9	2.03	0.18
	TT/TT	83	80.7	3.81	9.9×10^{-3}	72.3	5.22	1.6×10^{-3}	63.9	4.42	5.4×10^{-3}

OR - odds ratio

Supplementary Table 9. Comparison of rs12979860 and ss469415590 genotypes for predicting response to pegIFN-α/RBV treatment in European-Americans from HALT-C study. Response is evaluated as 20-week, end-of-treatment and sustained virological response (SVR) to pegIFN-α/RBV treatment in European-American (n=741) participants enrolled in the HALT-C Trial. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Variant	Genotype	N	20-week Response			End-of-Treatment Response			SVR		
			%	OR	p-value	%	OR	p-value	%	OR	*P* value
rs12979860	TT	129	15.5	Ref.		11.6	Ref.		7.0	Ref.	
	CT	425	29.6	2.30	$1.7x10^{-3}$	26.6	2.75	$6.2x10^{-4}$	14.6	2.28	0.027
	CC	187	66.8	10.99	$1.0 x10^{-16}$	58.3	10.61	$3.8 x10^{-14}$	33.7	6.77	$4.4 x10^{-7}$
ss469415590	ΔG/ΔG	133	15.8	Ref.		12.0	Ref.		6.8	Ref.	
	ΔG/TT	420	29.8	2.26	$1.8x10^{-3}$	26.7	2.66	7.0x10-^{4}	14.8	2.39	0.019
	TT/TT	188	66.5	10.58	$8.9 x10^{-17}$	58.0	10.09	$3.4 x10^{-14}$	33.5	6.94	$3.0 x10^{-7}$

OR - odds ratio

Supplementary Table 10. Association of rs12979860 and ss469415590 with spontaneous clearance of HCV infection among European-American injection drug users in the Urban Health Study (UHS). Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860 (n=557).

Variant	Genotype	Chronic N	%	Clear N	%	OR	*P* value	AUC *P* value
rs12979860	TT	41	10.4	6	3.7	Ref.		0.64
	CT	182	46.1	41	25.3	1.54	0.36	
	CC	172	43.5	115	71.0	4.57	$8.1x10^{-4}$	
	Total	395		162				
ss469415590	ΔG/ΔG	44	11.1	7	4.3	Ref.		0.64
	ΔG/TT	175	44.3	39	24.1	1.40	0.45	
	TT/TT	176	44.6	116	71.6	4.14	$8.1x10^{-4}$	
	Total	395		162				

OR - odds ratio; AUC - area under the receiver operating characteristic curve

Supplementary Table 11. Analysis of genetic variants based on Sanger sequencing and TaqMan genotyping of HapMap samples

variants	SNP ID	alleles	annotation	Allele 1	Allele 1 frequencies in HapMap samples		
					CEU	YRI	CHB/JPT
IFNL4, upstream	rs7248668	A/G	4,602 bp upstream of translation start	A	0.183	0.025	0.056
IFNL4, upstream	rs8099917	T/G	3,946 bp upstream of translation start	G	0.183	0.025	0.061
IFNL4, upstream	rs8109886	A/C	3,543 bp upstream of translation start	A	0.422	0.864	0.073
IFNL4, upstream	rs10853727	T/C	1,244 bp upstream of translation start	C	0.108	0.108	0.006
IFNL4, 5'UTR	rs4803222	G/C	5'UTR, 134 bp upstream of translation start	C	0.305	0.305	0.067
IFNL4, ex1	rs150891559	T/C	Ala11Ala	T	0.017	0	0
IFNL4, ex1	rs73555604	T/C	Cys17Tyr	T	0.025	0.275	0
IFNL4, ex1	ss469415590	ΔG/TT	Ala22/no IFNL4	ΔG	0.317	0.767	0.067
IFNL4, ex1	rs4803221	G/C	Ser30Ser	G	0.186	0.192	0.056
IFNL4, intron1	rs12979860	C/T	intron	T	0.317	0.700	0.067
IFNL4, intron1	rs181637919	C/T	intron	T	0.008	0.017	0
IFNL4, ex2	rs142981501	C/G	Pro60Arg	C	0	0.042	0
IFNL4, ex2	rs117648444	T/C	Pro70Ser	T	0.117	0.108	0.001
IFNL4, intron 3	rs111531283	T/G	intron	G	0.300	0.300	0
IFNL4, intron 3	rs143958949	C/A	intron	A	0	0.005	0
IFNL4, intron 4	rs77811741	G/T	intron	T	0	0.025	0
IFNL4, ex5	rs12971396	C/G	Ser149Ser	C	0.195	0.192	0.062
IFNL4, ex5	rs137902769	T/A	Ser175Ser	A	0.025	0.150	0
IFNL4, 3'UTR	ss539198934	T/C	3'UTR	C	0.033	0.117	0
IFNL3, 5'UTR	rs28416813	G/C	5'UTR	G	0.325	0.683	0.063
IFNL3, ex2	rs8103142	C/T	Lys70Arg	C	0.325	0.711	0

Supplementary Table 12. Analysis of pair-wise linkage disequilibrium (r^2) within *IFNL3/IFNL4* region in HapMap samples. The samples are from Europeans (CEU), West-Africans (YRI) and Asians - Chinese (CHB) and Japanese (JPT). Monomorphic markers are highlighted in light-gray, only marker with MAF>2.5% at least in one population are presented.

HapMap, CEU N=90 Markers	rs8099917	rs8109886	rs10853727	rs4803222	rs73555604	ss469415590	rs4803221	rs12979860	rs142981501	rs117648444	rs111531283	rs12971396	rs137902769	ss539198934	rs28416813	rs8103142
rs8099917																
rs8109886	0.320															
rs10853727	0.027	0.158														
rs4803222	0.522	0.609	0.282													
rs73555604	0.004	0.036	0.002	0.008												
ss469415590	0.420	0.640	0.262	0.885	0.055											
rs4803221	0.891	0.329	0.028	0.522	0.004	0.435										
rs12979860	0.420	0.640	0.262	0.885	0.055	0.923	0.435									
rs142981501																
rs117648444	0.011	0.173	0.767	0.246	0.002	0.227	0.005	0.227								
rs111531283	0.456	0.531	0.283	0.922	0.011	0.780	0.454	0.780		0.247						
rs12971396	0.946	0.394	0.03	0.551	0.005	0.463	0.946	0.463		0.009	0.483					
rs137902769	0.0	0.0	0.0	0.0	0.10	0.0	0.0	0.0		0.0	0.011	0.0				
ss539198934	0.154	0.049	0.004	0.08	0.0	0.074	0.153	0.074		0.006	0.08	0.145	0.001			
rs28416813	0.466	0.666	0.252	0.924	0.053	0.962	0.482	0.962		0.218	0.51	0.510	0.0	0.072		
rs8103142	0.466	0.666	0.197	0.850	0.053	0.889	0.482	0.889		0.168	0.51	0.510	0.0	0.072	0.925	

HapMap, YRI N=90 Markers	rs8099917	rs8109886	rs10853727	rs4803222	rs73555604	ss469415590	rs4803221	rs12979860	rs142981501	rs117648444	rs111531283	rs12971396	rs137902769	ss539198934	rs28416813	rs8103142
rs8099917																
rs8109886	0.004															
rs10853727	0.003	0.019														
rs4803222	0.059	0.069	0.282													
rs73555604	0.010	0.058	0.046	0.170												
ss469415590	0.008	0.504	0.037	0.130	0.115											
rs4803221	0.108	0.036	0.029	0.551	0.090	0.072										
rs12979860	0.011	0.357	0.052	0.178	0.163	0.710	0.102									
rs142981501	0.001	0.007	0.005	0.101	0.016	0.013	0.100	0.019								
rs117648444	0.003	0.018	1.0	0.282	0.042	0.034	0.026	0.052	0.005							
rs111531283	0.060	0.066	0.283	1.00	0.163	0.130	0.553	0.184	0.101	0.283						
rs12971396	0.108	0.036	0.029	0.551	0.090	0.072	1.000	0.102	0.100	0.026	0.553					
rs137902769	0.004	0.023	0.021	0.069	0.406	0.054	0.042	0.076	0.007	0.021	0.076	0.042				
ss539198934	0.003	0.019	0.016	0.059	0.348	0.04	0.031	0.057	0.006	0.016	0.057	0.031	0.748			
rs28416813	0.011	0.276	0.030	0.142	0.172	0.628	0.101	0.880	0.019	0.03	0.150	0.101	0.079	0.059		
rs8103142	0.011	0.301	0.023	0.133	0.158	0.610	0.093	0.800	0.018	0.023	0.133	0.093	0.073	0.055	0.920	

HapMap, CHB/JPT N=90 Markers	rs8099917	rs8109886	rs10853727	rs4803222	rs73555604	ss469415590	rs4803221	rs12979860	rs142981501	rs117648444	rs111531283	rs12971396	rs137902769	ss539198934	rs28416813	rs8103142
rs8099917																
rs8109886	0.004															
rs10853727																
rs4803222	0.059	0.069														
rs73555604																
ss469415590	0.008	0.504		0.130												
rs4803221	0.108	0.036		0.551		0.072										
rs12979860	0.011	0.357		0.178		0.710	0.102									
rs142981501																
rs117648444																
rs111531283	0.060	0.066		1.00		0.130	0.553	0.184								
rs12971396	0.108	0.036		0.551		0.072	1.000	0.102			0.553					
rs137902769				0.069		0.054	0.042	0.076			0.076					
ss539198934				0.059		0.04	0.031	0.057			0.057					
rs28416813	0.011	0.276		0.142		0.628	0.101	0.880			0.150	0.101				
rs8103142	0.011	0.301		0.133		0.610	0.093	0.800			0.133	0.093			0.920	

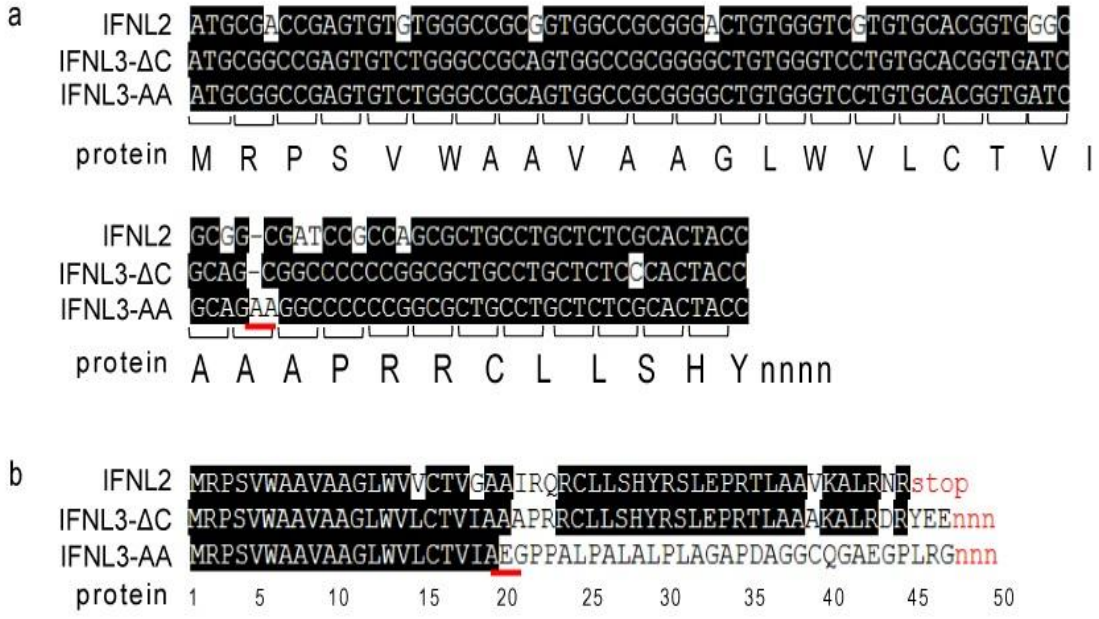
Supplementary Table 13. Haplotype analysis in the *IFNL3/IFNL4* region in HapMap samples and participants of Virahep-C study. Sustained virological response (SVR) was used as outcome of pegIFN-α/RBV therapy treatment (responders/non-responders). Bold underlined are markers included in the haplotype analysis of SVR by Smith et. al, 2011., * - 8 markers used in the final haplotype analysis in HapMap and Virahep-C samples. In yellow – the unfavorable haplotypes based on studies in Europeans and Asians, and their extrapolation in Africans; in grey – a common favorable haplotype shared by all populations. Highlighted bold underlined are non-synonymous variants within *IFNL4* found on the background of unfavorable haplotypes. Genotypes of *IFNL4* markers were determined by Sanger sequencing; no other coding variants were identified in these samples. EA- European-Americans; AA- African-Americans. Haplotype frequencies in all groups (HapMap and responders/non-responders) are indicated as %.

location	IFNL3			IFNL4									upstream of IFNL4				% in HapMap	Based on Smith et al, Genome Medicine, 2011, Table 2 n=819				Virahep-C EA: n=93 resp; n=85 non-resp; AA: n=43 resp; n=112 non-resp		
	intron 2	Lys70Arg	5'UTR	Ser175Ser, ex5	Ser149Ser, ex5	Pro70Ser, ex2	Arg60Pro, ex2	intron 1	Ser30Ser, ex1	Indel, ex1	Cys17Tyr, ex1	5'UTR	intergenic	intergenic	intergenic	intergenic								
variant	**rs11881222**	**rs8103142**	rs28416813	rs137902769	rs12971396*	rs11764844*	rs142981501*	**rs12979860***	**rs4803221 ***	ss469415590*	rs73555604*	rs4803222	**rs10853727**	**rs8109886**	**rs8099917***	**rs7248668**		resp, %	non-resp, %	p-value	OR, 95%CI	resp, %	non-resp, %	p-value
CEU n=90	G	C	G	T	**C**	C	G	T	**G**	**ΔG**	C	C	T	A	G	A	15.3	16.5	30.4	9.51E-11	2.20(1.72-2.80)	14.5	24.1	ref
	G	C	G	T	G	**T**	G	T	C	**ΔG**	C	C	**C**	A	T	G	9.3	10.1	10.6	0.77	1.04(0.75-1.43)	5.9	13.5	0.43
	A/G	C	G	T	G	C	G	T	C	**ΔG**	**T**	G	T	A	T	G	1.7	1.1	1.6	0.39	1.49 (0.62-3.56)	4.3	2.9	0.20
	A	T	C	T	G	C	G	C	C	TT	C	G	T	C	T	G	57.3	48.4	39.6	4.00E-04	0.70(0.57-0.85)	74.7	58.2	0.011
	A	T	C	T	G	C	G	C	C	TT	C	G	T	A	T	G	9.7	12.3	7.8	2.40E-03	0.60(0.43-0.83)			
YRI n=90	A	C	G	T/A	G	C	G	T	C	**ΔG**	**T**	G	T	A	T	G	24.8					14.0	22.3	
	G	C	G	T	**C**	C	G	T	**G**	**ΔG**	C	C	T	A	T	G	15.2					5.8	8.0	
	G	C	G	T	G	**T**	G	T	C	**ΔG**	C	C	**C**	A	T	G	10.0					9.3	4.9	
	A/G	C	G	T	G	C	G	T	C	**ΔG**	C	G	T	A	T	G	11.6					9.3	12.9	
	A	**T**	C	T	G	C	G	**C**	C	**ΔG**	C	G	T	A	T	G	6.7					1.1	7.6	
	G	C	G	T	**C**	C	**C**	T	**G**	**ΔG**	C	C	T	A	T	G	3.1					5.8	5.4	
	G	C	G	T	**C**	C	G	T	**G**	**ΔG**	C	C	T	A	G	A	2.5					9.3	6.7	
	A	C	G	T	G	C	**C**	T	C	**ΔG**	**T**	G	T	A	T	G	1.0					-	-	
	A	T	C	T	G	C	G	C	C	TT	C	G	T	C	T	G	12.5							
	A	T	C	T	G	C	G	C	C	TT	C	G	T	A	T	G	8.3					44.2	31.3	
	A	C	C	T	G	C	G	C	C	TT	C	G	T	A	T	G	1.7							
CHB, JPT, n=90	G	C	G	T	C	C	G	T	G	**ΔG**	C	C	T	A	G	A	5.5							
	A	T	C	T	G	C	G	C	C	TT	C	G	T	C	T	G	92.8							

Supplementary Fig. 1. DNA and protein sequence analysis of the regions upstream of the *IFNL2* and *IFNL3* genes. Identical nucleotides and amino acids are shaded in black. The location of ss469415590 is marked by a red bar.

a. Alignment of DNA sequences shows multiple mismatches between regions upstream of *IFNL2* and *IFNL3* genes. Protein translation is based on sequences upstream of *IFNL3* gene; M (Met) marks first amino acid of predicted proteins.

b. The region upstream of *IFNL2* is predicted to generate a protein fragment of 49 aa invariably terminated by a stop codon and expected to be degraded by nonsense-mediated decay. The region upstream of *IFNL3* gene is predicted to produce several full-length protein isoforms depending on alleles of ss469415590, ΔG and TT (ΔC and AA based on complementary DNA strand). The open reading frames and the identity of these proteins depend on ss469415590, a frame-shift variant at amino acid 22.



Supplementary Fig. 2. Sequence of *IFNL4* gene and genetic cariants found to be polymorphic at least in one population by sequencing of 270 HapMap samples (CEU, YRI and CHB/JPT).

Transcription start site

Gttgcccaggtggagacggctctggacgcctcccaggggacagtggacggcagcacctgctgcagcacgaggcacagagg
rs4803222
gtgcactgcagacaggagtgagggcagaggccaaggcgaggaggggccggctcccactctct**[C/G]**tcccactgtgtg
tgctgtgccttcacgctccgagcattgccttccctgggatcctaacccaaggcgggggctggacgcgctggaccctctc
tttggcttccctgacgtctctcgcctgctgcagaagcagag

Exon 1 rs150891559,Ala11Ala rs73555604,Cys17Tyr ss469415590
ATGCGGCCGAGTGTCTGGGCCGCAGTGGCC**GC[G/A]**GGGCTGTGGGTCCTG**T[G/A]C**ACGGTGATCGCA**G[AA/-C]G**GCCC
rs4803221,Ser30Ser
CCCGGCGCTGCCTGCT**CTC[G/C]**CACTACCGCTCGCTGGAGCCCCGGACGCTGGCGGCTGCCAAGGCGCTGAGGGACCGCTAC

Intron 1 rs117436747
gtaagtcaccgcccagcccctgtgccccctgggaccctggccccacc**[G/A]**ggttcccatacacccgttcctgtcccaaggg
gtcctgcgtcctagcgcccagcaggcgcctctcctatgtcagcgcccacaattcccaccacgagacccccgcagtccccgtcg
tcagcgcgaacgcaggctcagggtcaatcacagaagggagccctgccgggaggactcggctccaggtcggggcgaggggctttt
rs12979860
gctgggggagcgcggagtgcaattcaaccctggttc**[G/A]**cgccttcggggagctccctggttcagtacacgacaggcacga
rs181637919
c**[C/T]**gtgcgctgccagtacccatccacgtccaggaatcccagactgtgcagaggttaggggccctggcgaggggggcctagc
cgtatgcgataagcgccgcttgtcccgcag

Exon 2 (alternative) **rs142981501,Arg60Pro** **rs117648444,Pro70Ser**
GAGGAAGAGGCGCTGAGCTGGGGGCAG**C[G/C]C**AACTGCTCCTTCCGCCCCAGGAGGGAT**[C/T]CT**CCGCGGCCATCC
Intron 2
gtgaggcccgggagtgggcgggagaggcatggcccgggcgcggcccgctctaacgccctctcgtccccgcag

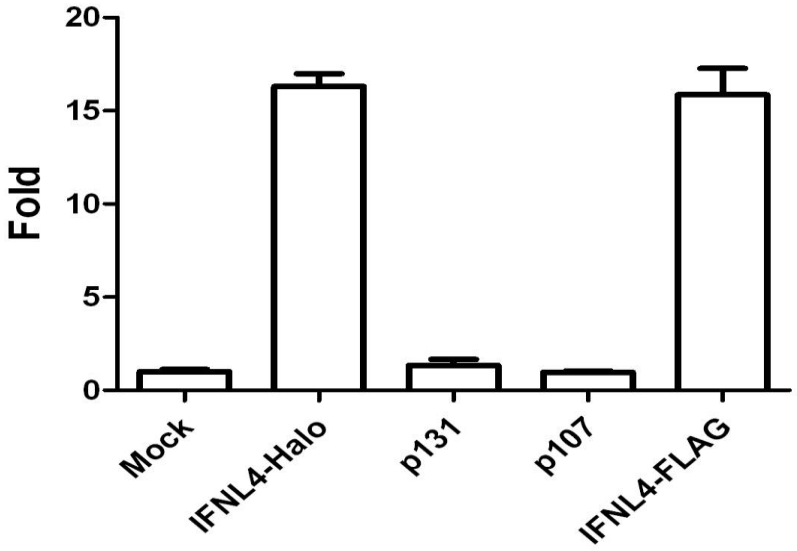
Exon 3(alternative)
TCCTGCGCTCGGCTCCGCCACGTGGCCCGGGGCATCGCGGACGCCCAGGCAGTGCTCAGCGGCCTGCACCGCTCGGAGCTGCT
CCCCGGCGCCGGCCCGATCCTGGAGCTGCTGGCGGCCGCGGGGAGGGATGTGGCGGCCTGC
Intron 3
rs111531283
gtgag**t**gacggccgcgccccgccgcccctc**[T/G]**cccccgccagcttctctgcatcctcaggcccacggcgagccccagcgc
rs143958949
tttgccaatctgtcctgcttagcggaaaaacccatccagac**[C/A]**ggagtcgggtcctctgggtgtcctgaaatccgggctc
gagtctgcggctgggagggccacgggcagatgcagagaggggcttcgtccttcgccttttccatttgcctcatgtcccacctc
cag

Exon 4
CTTGAGCTGGCACGGCCAGGCTCCTCCAGGAAGGTCCCCGGGGCCCAGAAGAGGCGTCACAAACCCCGGAGAGCG

Intron 4 **rs77811741**
gtgagtgcaacaggcaatacagg**[G/T]**ttagcccgcagggaggaccaggcgaggctgacaaggacgggactgaggctgcgag
cagcgggactggaggggattccgggggcgggggaagagcctggcttagccccgctgccctccctccctggctccag

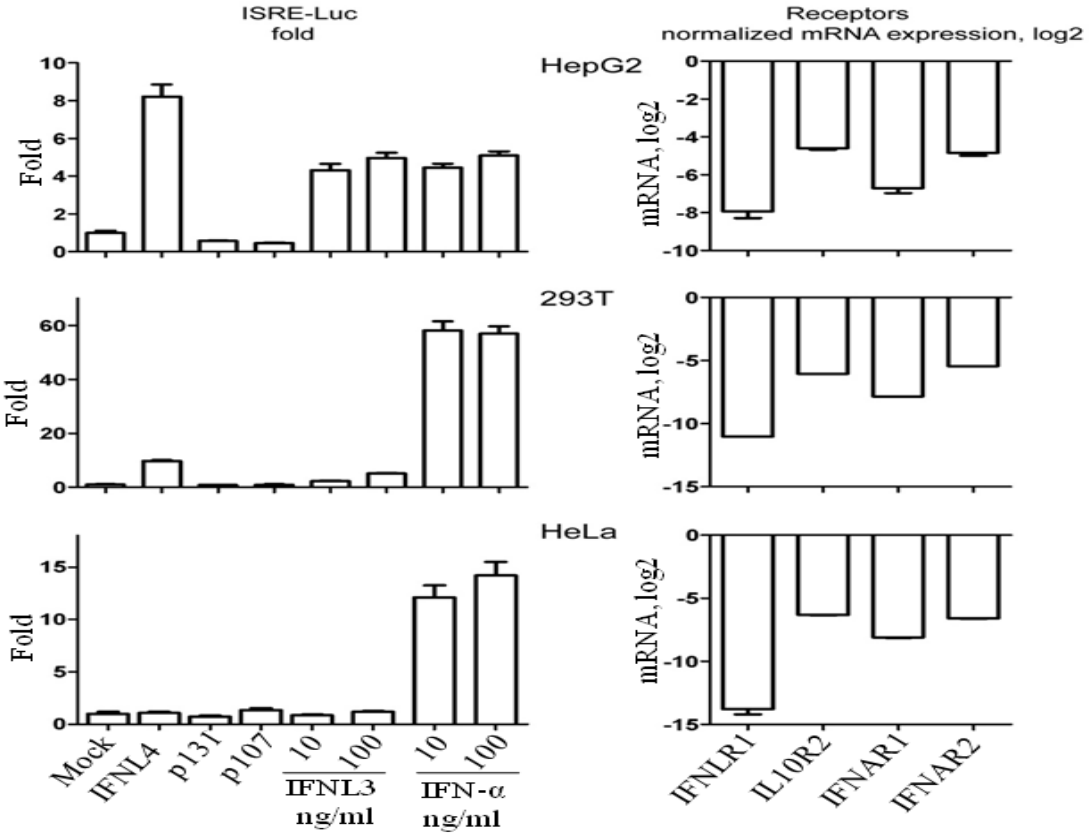
Exon 5
rs12971396,Ser149Ser
GAC**TC[G/C]**CCTCGGTGCCGCAAAGCCAGCGTGGTCTTCAACCTCCTGCGCCTGCTCACGTGGGAG**CTC**CGGCTGGCT
rs137902769,Ser175Ser ss539198934
GCACAC**TC[T/A]**GGGCCTTGCCTC**TGA**ccccgccccctc[T/C]ggcagcacggaaacctccacgccattggctgccgaaag

Supplementary Fig. 3. Activation of the Interferon Stimulated Response Element (ISRE) - Luc reporter in HepG2 cells transiently co-transfected with IFNL4 expression constructs carrying either the Halo-tag or a FLAG-tag. Alternatively, cells were transfected with expression constructs for non-functional forms p131 and p107, or an empty vector (mock). Similar activation of ISRE-Luc reporter is induced by IFNL4 with both protein tags, while mock, p131 and p107 did not induce activation. The results represent mean values of 8 biological replicates, with standard errors.



Nature Genetics: doi:10.1038/ng.2521

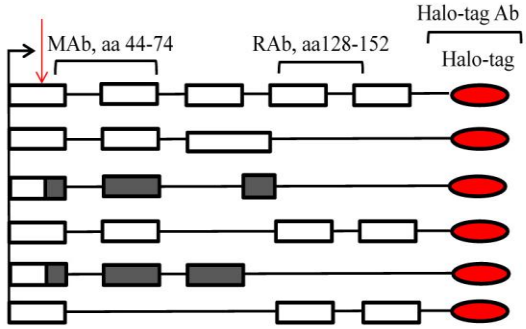
Supplementary Fig. 4. Activation of the interferon-stimulated response element (ISRE)-Luc reporter and mRNA expression of receptors for type–I IFNs (*IFNAR1* and *IFNAR2*) and type–III IFNs, *IFNLR1* (*IL28R1)* and *IL10R2*, in human cell lines HepG2, 293T and HeLa. Activation of ISRE-Luc reporter was analyzed after transient co-transfection with IFNL4, p131 or p107 expression constructs or after treatment with 10 and 100 ng/ml of IFN-α and IFNL3. Results represent mean values of 8 biological replicates, with standard errors. mRNA expression of interferon receptors was evaluated by qRT-PCR assays in non-treated cells and normalized to expression of four endogenous controls; less negative values represent higher mRNA expression. Expression of *IFNAR1*, *IFNAR2* and *IL10R2* was found at comparable levels in all cell lines tested. However, these cell lines differ by the level of expression of *IFNLR1*– the highest expression was detected in HepG2 cells, compared to which ~10 fold lower expression was detected in 293T cells and ~50 fold lower expression was detected in HeLa cell line.



Nature Genetics: doi:10.1038/ng.2521

Supplementary Fig. 5. Outline of expression constructs for the 6 protein isoforms. Epitopes recognized by the anti-IFNL4 mouse and rabbit monoclonal antibodies and the Halo-tag ab are indicated. The mouse monoclonal antibody recognizes IFNL4 (p179) and two non-functional proteins, p170 and p131; while the rabbit monoclonal antibody recognizes IFNL4 (p179) and two non-functional proteins, p131 and p107. All Halo-tag protein isoforms are recognized with the Halo-tag antibody in cell lysates of transiently transfected hepatoma HepG2 cells.

a

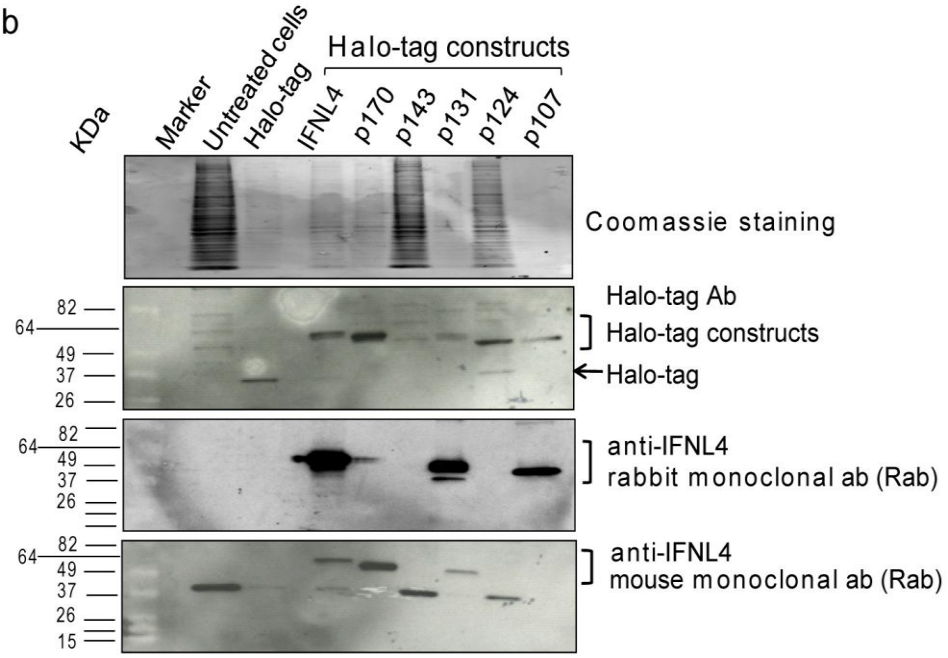


protein	Halo-tag Ab	IFNL4 -ab	
		MAb	RAb
IFNL4 p179 aa	Yes	**Yes**	**Yes**
p170 aa	Yes	**Yes**	No
p143 aa	Yes	No	No
p131aa	Yes	**Yes**	**Yes**
p124aa	Yes	No	No
p107aa	Yes	No	**Yes**

- transcripts with ΔG allele of ss469415590 (frame-shift), encode 4 protein isoforms, p179 (IFNL4) and 3 non-functional forms

- transcripts with TT allele of ss469415590, encode 2 protein isoforms unrelated to IFNL4

b



Nature Genetics: doi:10.1038/ng.2521

Supplementary Fig. 6. Analysis of IFNL4 expression in HepG2 cells and primary human hepatocytes.

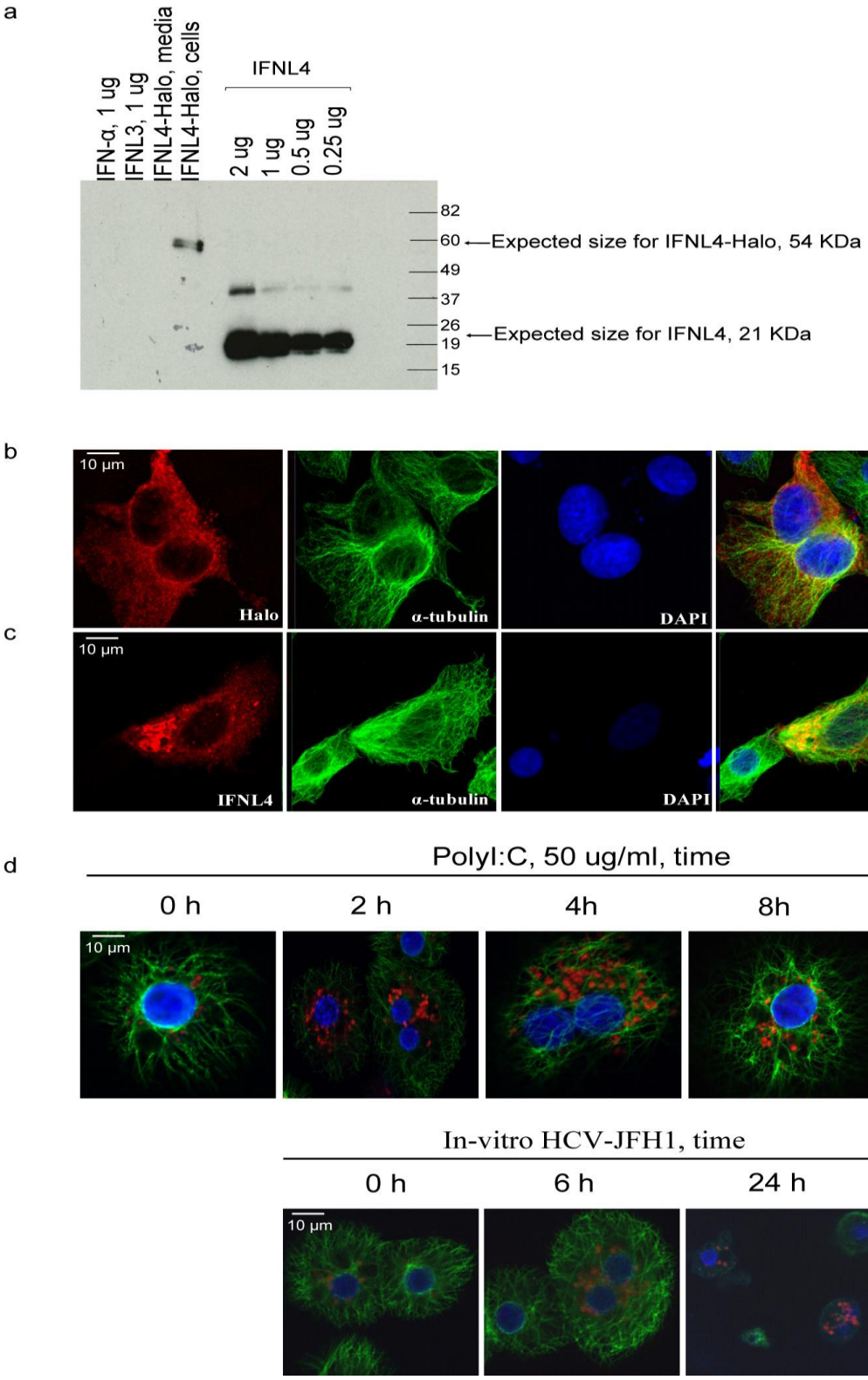
a. Western blot analysis of recombinant purified proteins (IFN-α, IFNL3 and IFNL4) and lysates of HepG2 cells transfected with IFNL4-Halo expression construct. Expression is detected with the mouse monoclonal anti-IFNL4 antibody. IFNL3 and IFNL4 are recombinant purified proteins generated in the sfs9 baculoviral system.

b. Confocal imaging with an anti-Halo-tag antibody in HepG2 cells transiently transfected with IFNL4-Halo construct.

c. Confocal imaging with a mouse monoclonal anti-IFNL4 antibody in HepG2 cells transiently transfected with IFNL4-Halo construct.

d. Confocal imaging in primary human hepatocytes from a liver donor not infected with HCV, who was heterozygous for ss469415590 ΔG/TT genotype. The cells were treated with 50 ug/ml of PolyI:C (the same image as in Fig. 5, ΔG/TT sample) or *in-vitro* infected with JFH1-HCV. The detection is performed with a mouse monoclonal anti-IFNL4 antibody. Red – IFNL4; green – α-tubulin (cytoplasm), blue – nuclei. Weak IFNL4 expression was detected even in samples not infected with JFH1-HCV (0 h). There was massive cell death in cells infected with JFH1-HCV at 24 h, but the consistency of this observation should be tested in additional samples.

Nature Genetics: doi:10.1038/ng.2521



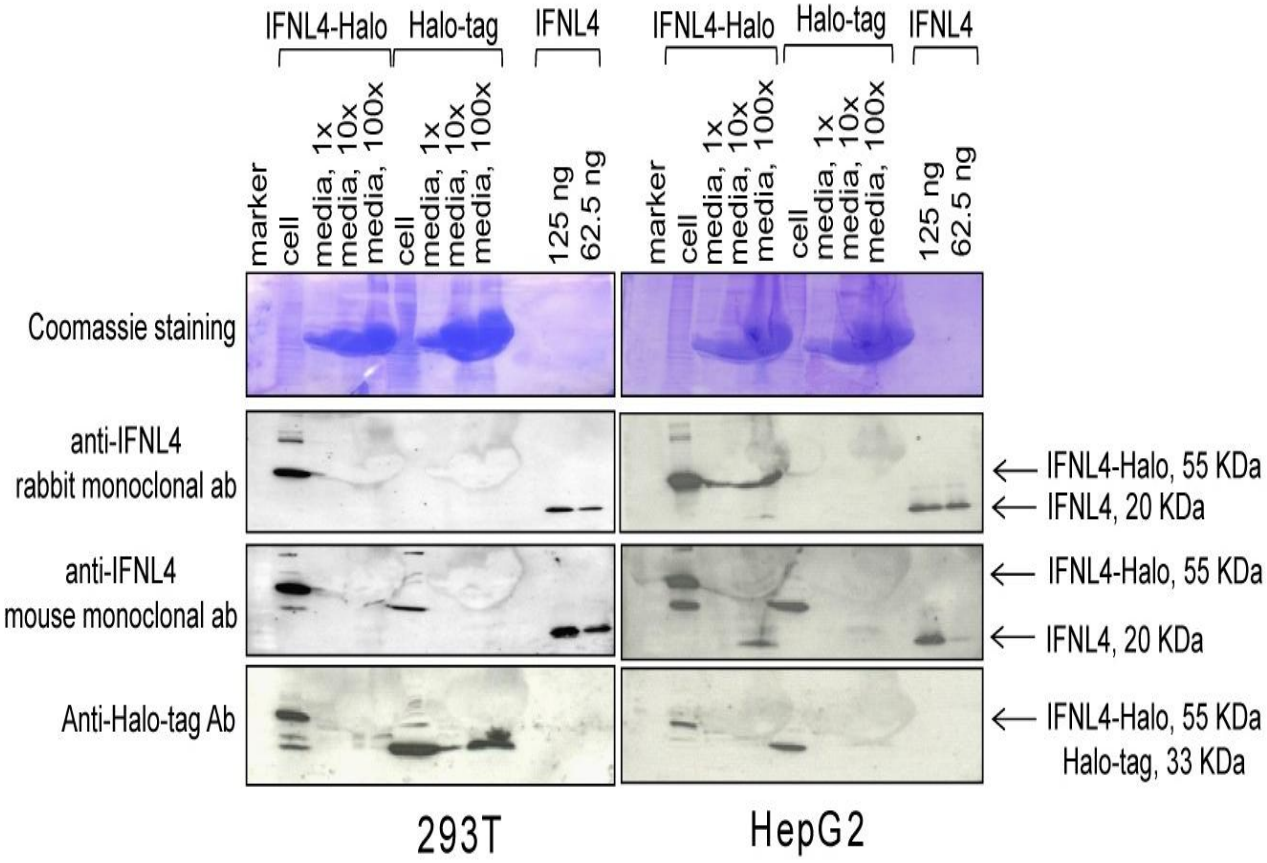
Nature Genetics: doi:10.1038/ng.2521

Supplementary Fig. 7. Analysis of IFNL4 expression in 293T and HepG2 cells transiently transfected with IFNL4-Halo construct or an empty Halo-tag vector.

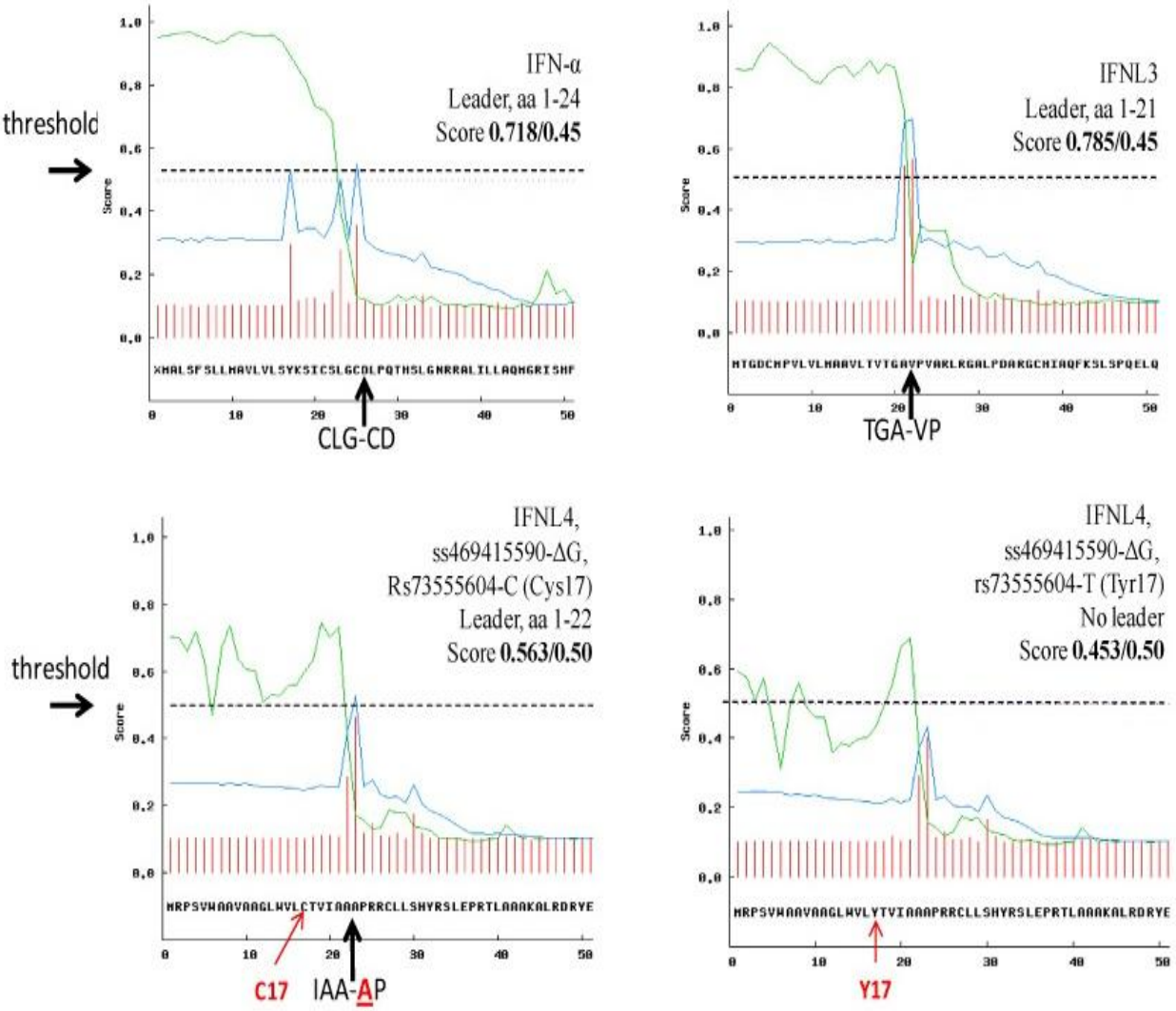
a. The cells and the media were collected 72 hours post-transfection, using one well of 6-well plate per condition. Cells were lysed in 300 ul of RIPA buffer and ~10 ul of each lysate was used per gel lane; cell media was used un-concentrated or concentrated 10x or 100x. Weak secreted IFNL4 expression was detected with the rabbit monoclonal anti-IFNL4 antibody in HepG2 cells, but not in 293T cells. Very weak or no secreted IFNL4 expression was detected with the mouse monoclonal anti-IFNL4 or a Halo-tag antibodies.

b. Prediction of leader peptides with SignalP4.0 server for secreted proteins IFN-α and IFNL3 and two allelic forms of IFNL4 that carry ss469415590-ΔG alleles but differ by rs73555604 (Cys17Tyr alleles). The weak leader peptide of IFNL4-Halo construct (with rs73555604, Cys17 variant) indicates that, compared to IFN-α and IFNL3, IFNL4 has lower secretability.

a.



b.


threshold
IFN-α
Leader, aa 1-24
Score 0.718/0.45
CLG-CD
IFNL3
Leader, aa 1-21
Score 0.785/0.45
TGA-VP
threshold
IFNL4,
ss469415590-ΔG,
Rs73555604-C (Cys17)
Leader, aa 1-22
Score 0.563/0.50
C17
IAA-AP
IFNL4,
ss469415590-ΔG,
rs73555604-T (Tyr17)
No leader
Score 0.453/0.50
Y17

Supplementary Fig. 8. Functional effects of IFNL4 over-expression in HepG2 cells

a. RNA-seq of HepG2 cells transfected with an empty vector (mock) or the IFNL4 expression construct. RNA-seq reads contributed by the IFNL4 expression construct are marked as 'IFNL4'. The results are presented as clusters of RNA-seq reads, with the number of reads (depth) corresponding to the level of mRNA expression. Within the 150-Kb region, only expression of *IFNL1* is induced by IFNL4 over-expression.

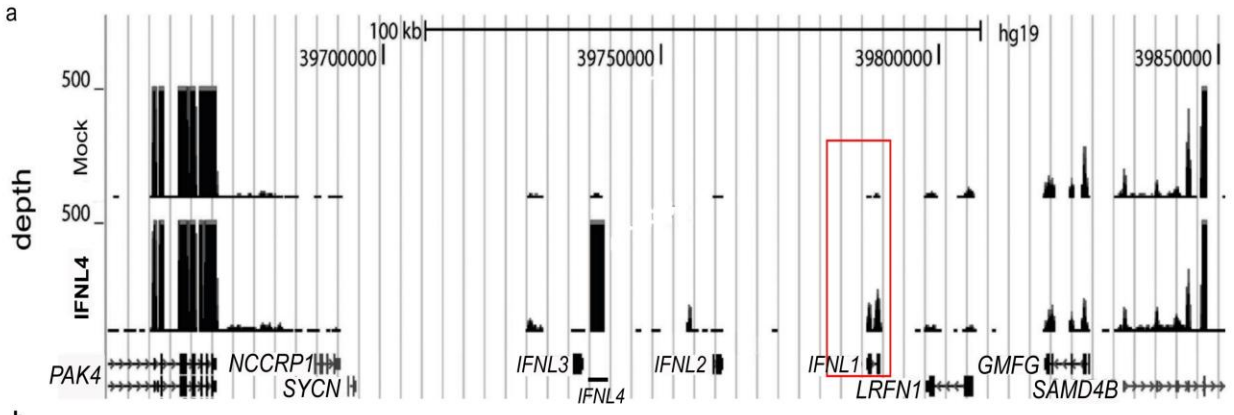
b. Top canonical pathways statistically significantly affected by IFNL4 over-expression in HepG2 cells, according to Ingenuity Pathway Analysis (IPA). The ratio indicates the number of transcripts affected by IFNL4 over-expression compared to the total number of transcripts in the corresponding pathways.

c. List of transcripts (from panel b) enriched in the corresponding pathways, Log2 ratio indicates the magnitude of expression change in IFNL4-transfected compared to mock-transfected samples.

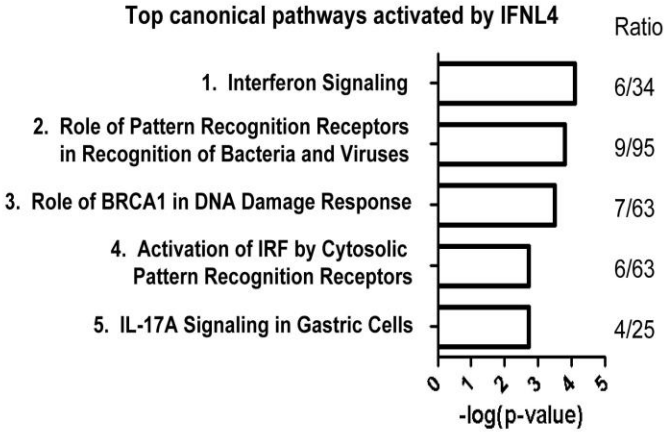
d. Quantitative reverse-transcriptase PCR (qRT-PCR) expression analysis of transcripts selected from the list (presented on panel c) and tested in HepG2 cells in specified conditions: untreated cells, mock-transfected cells or cells transfected with IFNL4, p131 and p107 expression constructs and/or treated with 10 ng/ml or purified recombinant IFN-α and IFNL3 proteins. Expression is presented on log 2 scale as ΔCt values, which correspond to Ct values (PCR cycle at detection) of target genes normalized by an average of Ct values of 4 endogenous controls. Less negative ΔCt values correspond to higher mRNA expression of target genes. The fold difference between any two samples can be roughly calculated as fold=$2^{\wedge(\Delta Ct\ sample1-\Delta Ct\ sample\ 2)}$. Equal amounts of DNase-treated RNA were used for all the assays. Results are presented as mean values of 3 or 4 biological replicates, with standard errors. Expression analysis was done with The Human Antiviral Response RT² Profiler™ PCR Array (Qiagen).

Nature Genetics: doi:10.1038/ng.2521

a



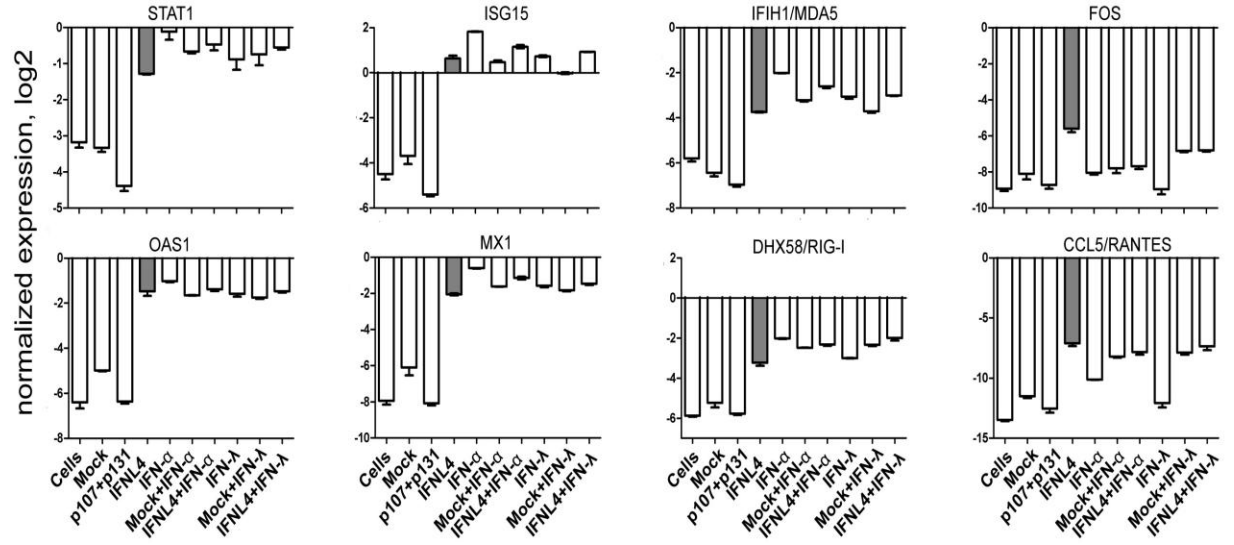
b



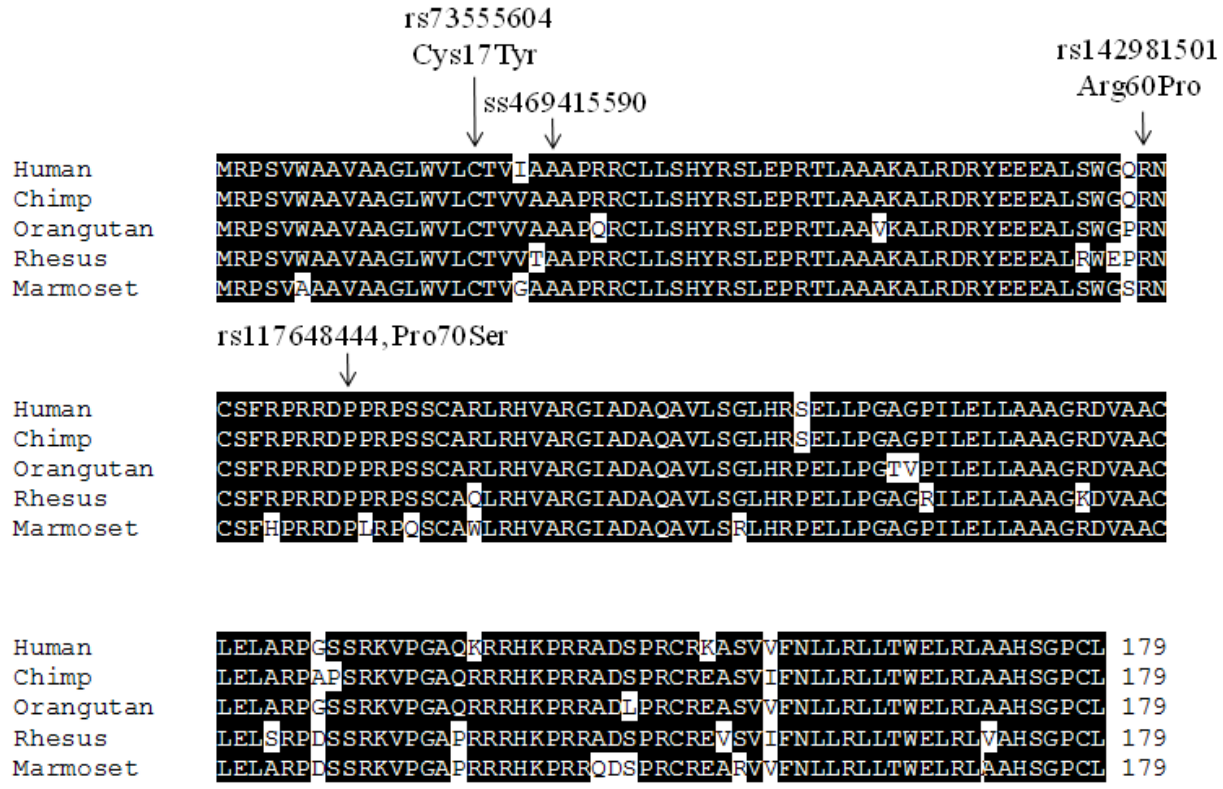
c

Symbol	Log2 Ratio	Pathway
IFITM1	3.83	1
IFIT1 (IFI-56)	3.53	1
IFIT3	3.51	1
MX1	2.80	1
OAS1	1.48	1, 2
STAT1	2.38	1, 3, 4
OAS2	4.48	2
TLR3	3.92	2
EIF2AK2 (PKR)	2.47	2
OAS3	2.29	2
PIK3C2A	1.44	2
IFIH1 (MDA5)	2.79	2, 4
DDX58 (RIG-I)	2.22	2, 4
CCL5 (RANTES)	2.66	2, 5
BRCA2	2.32	3
FANCB	2.16	3
BARD1	1.60	3
NBN	1.59	3
FANCM	1.54	3
RFC3	1.52	3
ZBP1	4.08	4
IFIT2 (IFI-54)	3.69	4
ISG15	2.24	4
CXCL11 (IP-9)	3.36	5
IL8	1.70	5
FOS	1.60	5

d



Supplementary Fig. 9. Prediction of IFNL4 protein based on genomic sequences of 45 species available in UCSC genome browser (genome.uscs.edu). IFNL4 protein is predicted only in the genomes of macaque (marmoset and rhesus), orangutan, chimpanzee and human. Identical amino acids are shaded in black and the positions of human non-synonymous genetic variants are indicated.



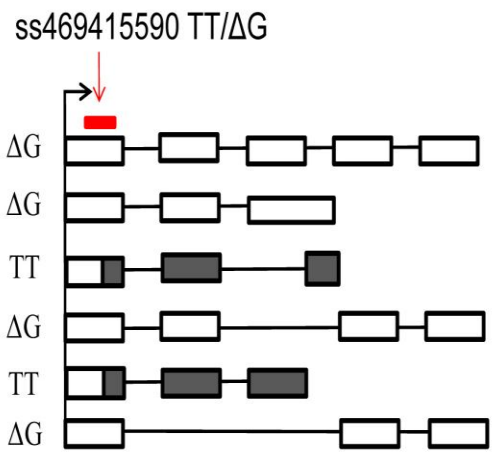
Supplementary Fig. 10. Outline of mRNA transcripts within *IFNL4* region and their detection by RNA-sequencing and an allele-specific expression assay.

a. Outline of all mRNA transcripts generating full-length protein isoforms. An allele-specific expression assay for ss469415590 TT/ΔG is located within the first exon of *IFNL4* and detects all transcripts with TT and ΔG alleles. Thus, this assay does not discriminate between the four transcripts with ΔG allele generating the functional IFNL4 (p179) protein and non-functional forms p170, p131 and p107. RNA-seq results show fragments per kilobase of exon per million fragments mapped (FPKM) values which roughly correspond to relative levels of expression of each of the transcripts detected by RNA-seq analysis of PolyI:C-stimulated primary human hepatocytes from a liver donor heterozygous for ss469415590 (Fig. 1). Of all the transcripts carrying the ΔG allele, the transcript for the non-functional p170 is the most abundant isoform at 2, 4, 8 and 24 hours of PolyI:C treatment. The functional IFNL4 (p179) transcript is the second common isoform at 4, 8 and 24 hours after PolyI:C treatment.

b. Outline of DNA/cDNA sequence and the primers used for an allele-specific genotyping and expression assay for ss469415590 TT/ΔG (corresponds to AA/-C in on a complementary strand). The full information about this assay is presented in **Supplementary Table 1**. The assay targets ss469415590 and avoids other variants in this amplicon. Since the assay is located within exonic sequence, it can be used both for genotyping and expression analysis, for which RNA has to be DNAse-treated to eliminate any possible signal from residual contaminating DNA.

Nature Genetics: doi:10.1038/ng.2521

a



RNA-seq, FPKM				
transcript	2 h	4 h	8 h	24 h
IFNL4 p179aa	0	96.2	31.7	18.5
p170aa	133.4	233.2	71.6	50.0
p143aa	1.3	0	0	1.5
p131aa	0	39.3	11.3	0
p124aa				
p107aa	71.7	0	0	8.9

- transcripts with ΔG allele of ss469415590 (frame-shift), encode 4 protein isoforms, p179 (IFNL4) and 3 non-functional forms

-transcripts with TT allele of ss469415590, encode 2 protein isoforms unrelated to IFNL4

b

forward primer

GCCTGCTGCAGAAGCAGAGATGCGGCCGAGTGTCTGGGCCGCAGTG

rs150891559 rs73555604 **ss469415590**

GCCGC[**G/c**]GGGCTGTGGGTCCTGT[**G/a**]CACGGTGATCGCAG**[AA/-C]**

rs4803221 reverse primer

GGCCCCCCGGCGCTGCCTGCTCTC[**G/c**]**CACTACCGCTCGCTGGAGC**

Nature Genetics: doi:10.1038/ng.2521

Supplementary Fig. 11. Quantitative reverse-transcriptase PCR (qRT-PCR) analysis of allele-specific expression of transcripts with ss469415590-TT and ΔG alleles and expression of *IFNL3 (IL28B)*, *IFNL1 (IL29)* and *PPIA* (endogenous control) in various samples. Expression is presented on log 2 scale as Ct values (PCR cycle at detection), with higher Ct values corresponding to lower mRNA expression. The fold difference between any two samples can be roughly calculated as fold=$2^{\wedge(\text{Ct sample1-Ct sample 2})}$. Equal amounts of DNase-treated RNA were used for all the assays. Graphs show activation of transcripts carrying both ss469415590 alleles (TT and ΔG). Transcripts carrying the risk ss469415590-ΔG allele generate the functional IFNL4/p179 protein, as well as non-functional proteins p170, p131, p107, while transcripts with the beneficial ss469415590-TT allele generate only non-functional proteins. All the samples were also genotyped for rs12979860 and ss469415590 alleles TT and ΔG correspond to rs12979860 alleles C and T. All primary human hepatocytes (PHH) are from liver donors not infected with HCV.

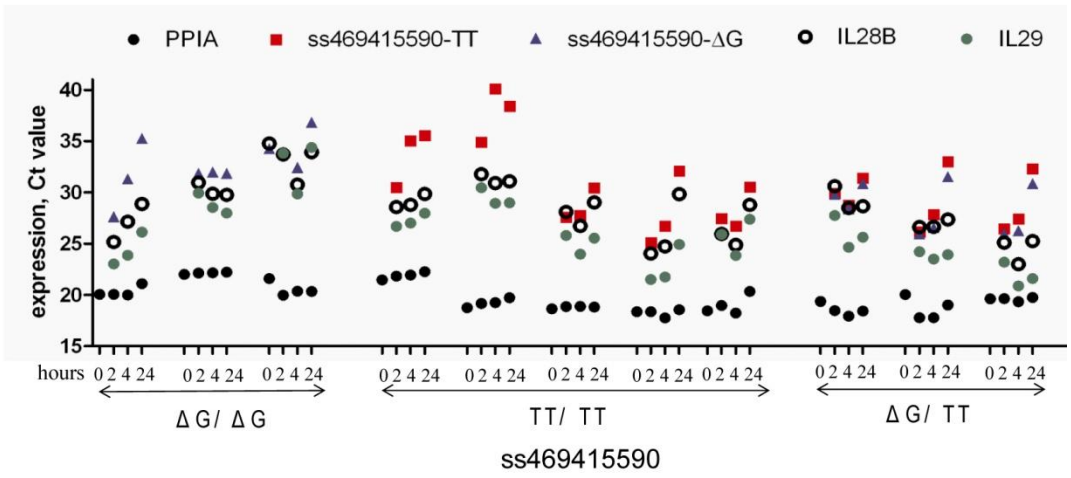
a. Expression analysis in PHH treated with 50 ug/ml of PolyI:C for 0, 2, 4 and 24 hours. PHH are from liver donors with ss469415590 ΔG/ΔG (n=3), TT/TT (n=5) and TT/ΔG (n=3) genotypes.

b. Expression analysis in PHH *in-vitro* infected with JFH1-HCV for 0, 6 and 24 hours. PHH are from liver donors with ss469415590 ΔG/ΔG (n=1) and TT/TT (n=3) genotypes.

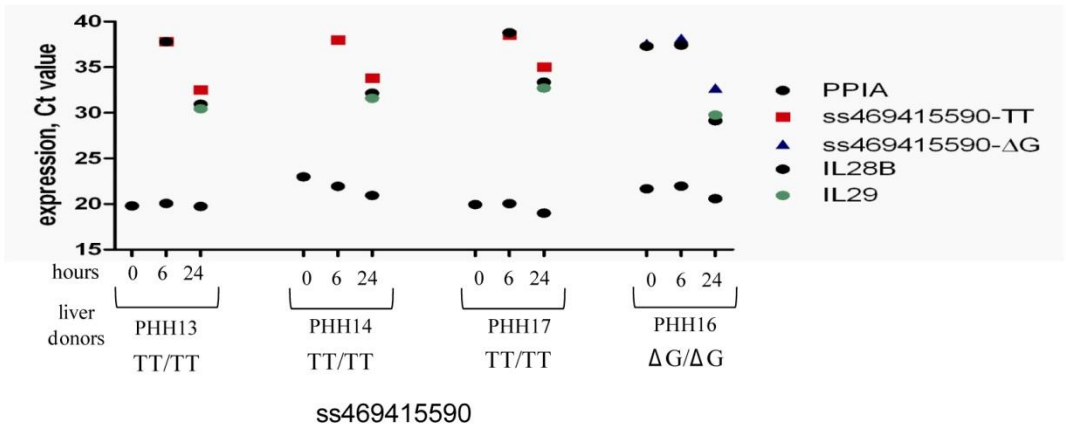
c. Expression analysis in PHH or HepG2 cells. PHH are from a liver donor with ss469415590 ΔG/ΔG genotype. PHH were untreated or treated with 100 ng/ml of IFN-α or IFN-λ. Only IFN-α treatment induced low *IFNL4* expression. HepG2 cells with ss469415590 TT/ΔG genotype were treated with 50 ug/ml of PolyI:C for 0, 1, 2, 4, 8 and 24 hours, but no expression of transcripts with either TT or ΔG alleles of ss469415590 was induced.

Nature Genetics: doi:10.1038/ng.2521

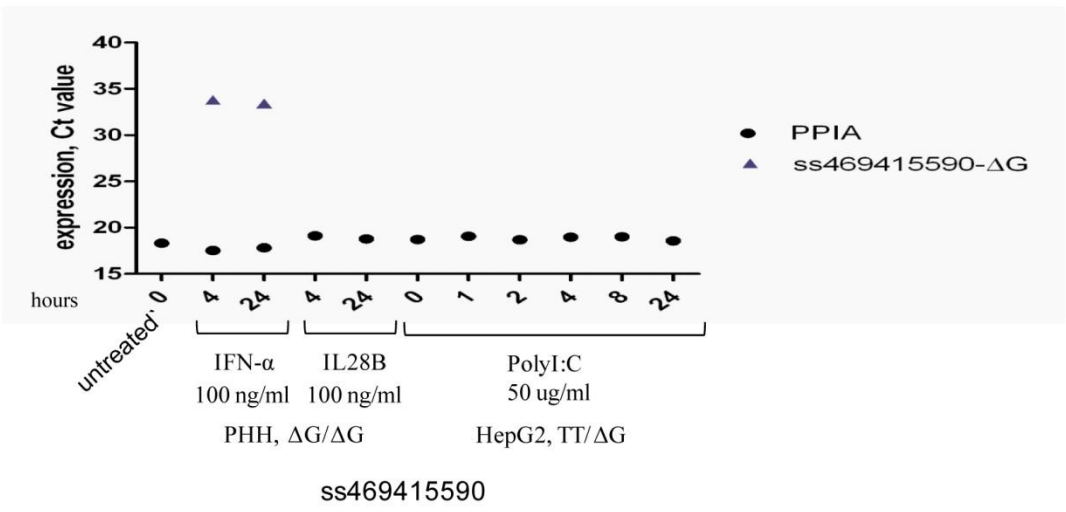
a



b



c



Nature Genetics: doi:10.1038/ng.2521

Supplementary Note - Clinical Studies

VirahepC

The Study of Viral Resistance to Antiviral Therapy of Chronic Hepatitis C (Virahep-C) was designed to compare response to treatment with pegylated IFN-α/ribavirin in African American patients with chronic hepatitis C to otherwise similar patients of European ancestry[1]. In Virahep-C, patients with HCV genotype 1 infection who had not undergone previous treatment for chronic hepatitis C received treatment with a standard regimen of pegylated IFN-alfa-2a (180 μg/week) plus ribavirin (1000-1200 mg/day) for up to 48 weeks. Ancestral designation was self-reported. Study end points included: decrease in HCV RNA levels between baseline and various treatment time points; week 24 response (absence of detectable HCV RNA in serum after 24 weeks of therapy); end-of-treatment response (absence of HCV RNA after 48 weeks of therapy); and sustained virologic response (SVR; absence of HCV RNA 24 weeks after treatment was stopped). The protocol was approved by the institutional review boards of the participating institutions and all patients gave informed written consent. Reported results from Virahep-C showed that African-American patients had lower rates of virologic response than European-American patients and that those differences were not explained by differences in patient characteristics, baseline HCV RNA levels or the amount of medication taken during the study[1].

HALT-C

The Hepatitis C Antiviral Long-term Treatment against Cirrhosis (HALT-C) Trial was a study of patients with advanced chronic hepatitis C who had failed previous interferon-based treatment[2,3]. At enrollment, HALT-C patients had an Ishak fibrosis score >3 by local assessment of liver biopsy, had a Child-Turcotte-Pugh score <7 and had no evidence of hepatocellular carcinoma. Final assessment of fibrosis stage was performed by a panel of hepatopathologists. Patients with other liver diseases, human immunodeficiency virus infection, active illicit drug use or current alcohol abuse were excluded. Ancestral designation was self-reported. During the lead-in phase of HALT-C, patients underwent retreatment with pegylated-interferon-alfa-2a (180 μg/week) plus ribavirin (1000-1200 mg/day). Subjects with undetectable HCV RNA at week 20 remained on combination treatment through week 48 and were followed until week 72. Subjects

Nature Genetics: doi:10.1038/ng.2521

with undetectable HCV RNA at weeks 48 and 72 were considered to have an SVR. Investigations of human genetics in the HALT-C Trial were conducted in those participants who provided (written) consent for genetic testing. The HALT-C Trial was approved by institutional review boards of the participating institutions.

Urban Health Study (UHS)

As previously described[4], UHS recruited IDUs from street settings in six inner-city San Francisco Bay area neighborhoods from 1986 through 2002, drawing serial cross-sectional samples every six months[5]. Individuals 18 years of age or older were eligible for enrollment if they had injected drugs within the past 30 days or previously enrolled in the UHS study. New participants were screened for visible signs of recent or chronic injection (i.e., venipuncture sites or scars). Interviews were conducted using standardized questionnaires and blood samples were collected from participants who provided written informed consent. The present study included unduplicated IDUs recruited between 1998 and 2000[6]. Participants who were positive for HCV antibody were divided into two groups based on their HCV RNA result: 'chronic' (positive for HCV RNA) or 'cleared' (negative for HCV RNA and positive for antibody). All subjects with cleared infection were included in the study and frequency matched to those with chronic infection (maximum 4:1) on the basis of self-reported ethnicity and age.

Information on demographic variables and other potential covariates were assessed through face-to-face personal interviews[6,7]. After being interviewed participants were counseled by trained staff on reducing infection risks and referred to appropriate medical and social services. Participants were not asked about treatment for HCV infection during 1998-2000, but when they were asked during 2001-2002, reports of antiviral treatment were rare[8]; thus it is very likely that the HCV seropositive, HCV RNA negative subjects in this study had recovered spontaneously. For the purpose of genetic investigations, ancestry was ascertained by self report; subjects who reported themselves to be White and not of Latino/Hispanic ethnicity are considered to be of 'European American' ancestry. Study procedures were approved by an Institutional Review Board of the National Cancer Institute and the Committee on Human Subjects Research at the University of California, San Francisco.

ALIVE

The AIDS Link to Intravenous experience (ALIVE) is an ongoing study of injection drug users enrolled in Baltimore, Maryland, from February 1988 through March 1989[9] .HCV infection was established by detection of HCV antibody (anti-HCV) by enzyme immunoassay (EIA) and recombinant immunoblot assay (RIBA [version 3.0]; Novartis). Individuals with cleared HCV infection had anti-HCV (as confirmed by RIBA) and undetectable HCV RNA in serum or plasma without having received any HCV therapy. Individuals with persistent infection had anti-HCV and HCV RNA in serum or plasma before receiving any HCV therapy. Written informed consent for genetic testing was obtained from all participants. The study was approved by the institutional review board at Johns Hopkins University.

References for Supplementary Note:

1. Conjeevaram, H.S. et al. Peginterferon and ribavirin treatment in African American and Caucasian American patients with hepatitis C genotype 1. *Gastroenterology* **131**, 470-7 (2006).
2. Di Bisceglie, A.M. et al. Prolonged therapy of advanced chronic hepatitis C with low-dose peginterferon. *N Engl J Med* **359**, 2429-41 (2008).
3. Lee, W.M. et al. Evolution of the HALT-C Trial: pegylated interferon as maintenance therapy for chronic hepatitis C in previous interferon nonresponders. *Control Clin Trials* **25**, 472-92 (2004).
4. 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**, 1843-7.
5. Watters, J.K., Bluthenthal, R.N. & Kral, A.H. HIV seroprevalence in injection drug users. *JAMA* **273**, 1178 (1995).
6. Tseng, F.C. et al. Seroprevalence of hepatitis C virus and hepatitis B virus among San Francisco injection drug users, 1998 to 2000. *Hepatology* **46**, 666-71 (2007).
7. Kral, A.H. et al. Sexual transmission of HIV-1 among injection drug users in San Francisco, USA: risk-factor analysis. *Lancet* **357**, 1397-401 (2001).
8. Seal, K.H. et al. Among injection drug users, interest is high, but access low to HCV antiviral therapy. Society of General Internal Medicine, 28th annual meeting, New Orleans, Louisiana, USA, May 11-14, 2005. *J Gen Intern Med* **20 Suppl 1**, 171 (2005).
9. 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).

Nature Genetics: doi:10.1038/ng.2521