Prokunina-Olsson *et al.*

A Genetic Variant Creating a Novel Interferon Analog (IFNAN) Protein is Associated with Impaired Clearance of Hepatitis C Virus

Supplementary Table 1. Description of novel transcripts and proteins identified upstream of *IL28B* gene

Transcript ID, NCBI accession number	Exons	mRNA	Protein	ss469415590 allele	Annotation
IFNAN, p179, JN806234	5	1636 bp	179 aa	ΔG	*Similarity to IL28B, full-length protein
p131, JN806225	4	1492 bp	131 aa	ΔG	Similarity to IL28B, full-length protein
p107, JN806226	3	1420 bp	107 aa	ΔG	Similarity to IL28B, 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. IFNAN (p179) has highest homology with IL28B – 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.

Supplementary Table 2. Allele frequencies for ss469415590 and GWAS markers rs12979860 and rs8099917 in HapMap populations reference panels. 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

Supplementary Table 3. 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 4. Characteristics of the Virahep-C and HALT-C participants

Characteristic	Virahep-C (n=351)		HALT-C (n=885)	
Age (median, IQR)	48.6	44.0-52.8	49	46.0-53.0
Race				
European-Americans (n, %)	182	51.9	741	83.7
African-Americans (n, %)	169	48.1	144	16.3
Male (n, %)	230	65.5	641	72.4
Ishak Fibrosis Score [b]				
2 (n, %)	98	27.9	68	7.7
3-4 (n, %)	104	29.6	495	55.9
5-6 (n, %)	25	7.1	321	36.3
HCV Genotype 1 (n, %)	351	100	794	89.7
HCV RNA level (log 10 IU) (median, IQR[a])	6.5	5.6-6.7	6.5	6.1- 6.8
Prior treatment:				
None	351	100		
pegIFN-α alone (n, %)	0	0	243	27.5
pegIFN-α/RBV (n, %)	0	0	642	72.5

IQR = interquartile range

Supplementary Table 5. Characteristics of the UHS and ALIVE participants

Characteristic	Urban Health Study (UHS)				ALIVE	
	European-Americans		African-Americans		African-Americans	
	Chronic	Cleared	Chronic	Cleared	Chronic	Cleared
N	395	162	350	109	590	81
Age, median (IQR)	42.0 (36 – 48)	39.5 (33 – 47)	46.0 (42 – 50)	46.0 (42 – 49)	40 (36 – 45)	40 (36 – 43)
Years injection drug use, median (IQR)	23.0 (16 – 29)	20.0 (11 – 29)	27.0 (21 – 32)	25.0 (19 – 31)	14 (9 – 19)	13 (8 – 19)
Male (%)	72.4	67.3	62.6	60.6	76	65
HIV-1-infected (%)	12.2	4.9	13.1	10.1	54	32
Chronic HBV (%)	3.3	4.3	2.9	9.2	3	7

IQR = interquartile range

Supplementary Table 6. Comparison of rs12979860 and ss469415590 genotypes for predicting the median decrease in HCV RNA load (log_{10} IU/ml) 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.

Race	Variant	Genotype	N	%	HCV RNA Decrease (Median, log_{10} IU/ml)	HCV RNA Difference*	p-value**
EA n=176	**rs12979860**	TT	19	10.8	1.14	Ref.	.
		CT	78	44.3	1.57	0.43	0.03
		CC	79	44.9	2.99	1.85	3.4E-07
	ss469415590	ΔG/ΔG	19	10.8	1.14	Ref.	.
		ΔG/TT	78	44.3	1.57	0.43	0.03
		TT/TT	79	44.9	2.99	1.85	3.4E-07
AA n=162	**rs12979860**	TT	53	32.7	1.05	Ref.	.
		CT	91	56.2	1.31	0.26	0.03
		CC	18	11.1	2.70	1.64	0.0001
	ss469415590	ΔG/ΔG	67	41.4	0.87	Ref.	.
		ΔG/TT	80	49.4	1.54	0.66	8.4E-05
		TT/TT	15	9.3	2.84	1.97	6.6E-07

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

Supplementary Table 7. Comparison of rs12979860 and ss469415590 genotypes for predicting 24-week, end-of-treatment and sustained virological response (SVR) to pegIFN-α/RBV treatment in European-American (EA) and African-American (AA) participants in Virahep-C. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Race	Variant	Genotype	N	24-week Response			End-of-Treatment Response			SVR		
				%	OR	p-value	%	OR	p-value	%	OR	p-value
EA n=182	**rs12979860**	TT	21	52.4	Ref.		33.3	Ref.		28.6	Ref.	
		CT	78	74.4	2.64	0.06	67.9	4.24	0.006	44.9	2.03	0.18
		CC	83	80.7	3.81	0.01	72.3	5.22	0.002	63.9	4.42	0.005
	ss469415590	ΔG/ΔG	21	52.4	Ref.		33.3	Ref.		28.6	Ref.	
		ΔG/TT	78	74.4	2.64	0.06	67.9	4.24	0.006	44.9	2.03	0.18
		TT/TT	83	80.7	3.81	0.01	72.3	5.22	0.002	63.9	4.42	0.005
AA n=169	**rs12979860**	TT	57	40.4	Ref.		36.8	Ref.		24.6	Ref.	
		CT	93	50.5	1.51	0.23	40.9	1.18	0.62	29.0	1.26	0.55
		CC	19	63.2	2.53	0.09	52.6	1.90	0.23	36.8	1.79	0.30
	ss469415590	ΔG/ΔG	71	35.2	Ref.		32.4	Ref.		21.1	Ref.	
		ΔG/TT	82	56.1	2.35	0.01	45.1	1.72	0.11	31.7	1.73	0.14
		TT/TT	16	68.8	4.05	0.02	56.3	2.68	0.08	43.8	2.90	0.07

OR = Odds Ratio

Supplementary Table 8. Comparison of rs12979860 and ss469415590 genotypes for predicting 20-weeks, end-of-treatment and sustained virological response (SVR) to pegIFN-α/RBV treatment in European-American (EA) and African-American (AA) participants enrolled in the HALT-C Trial. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Race	Variant	Genotype	N	20-week Response			End-of-Treatment Response			SVR		
				%	OR	p-value	%	OR	p-value	%	OR	p-value
EA n=741	**rs12979860**	TT	129	15.5	Ref.		11.6	Ref.		7.0	Ref.	
		CT	425	29.7	2.30	0.002	26.6	2.75	0.0006	14.6	2.28	0.03
		CC	187	66.8	10.99	1.0E-16	58.3	10.61	3.8E-14	33.7	6.77	4.4E-07
	ss469415590	ΔG/ΔG	133	15.8	Ref.		12.0	Ref.		6.8	Ref.	
		ΔG/TT	420	29.8	2.26	0.002	26.7	2.66	0.0007	14.8	2.39	0.02
		TT/TT	188	66.5	10.58	8.9E-17	58.0	10.09	3.4E-14	33.5	6.94	3.0E-07
AA n=144	**rs12979860**	TT	61	9.8	Ref.		8.2	Ref.		3.3	Ref.	
		CT	72	16.7	1.83	0.26	12.5	1.60	0.42	6.9	2.20	0.36
		CC	11	45.5	7.64	0.006	36.4	6.40	0.02	27.3	11.06	0.01
	ss469415590	ΔG/ΔG	68	8.8	Ref.		7.4	Ref.		2.9	Ref.	
		ΔG/TT	68	19.1	2.44	0.009	14.7	2.17	0.18	8.8	3.19	0.16
		TT/TT	8	50.0	10.33	0.005	37.5	7.56	0.02	25.0	11.0	0.03

OR = Odds Ratio

Supplementary Table 9. Association of the GWAS marker rs12979860 and ss469415590 with spontaneous clearance of hepatitis C virus infection among European-American and African-American injection drug users enrolled in the Urban Health Study (UHS). Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Variant	Genotype	Chronic	%	Clear	%	OR	p-value	AUC p-value
European-Americans, n=557								
rs12979860	TT	41	10.4	6	3.7	Ref.		0.64
	CT	182	46.1	41	25.3	1.5	0.36	
	CC	172	43.5	115	71	4.6	0.0008	
	Total	395		162				
ss469415590	ΔG/ΔG	44	11.1	7	4.3	Ref.		0.64
	ΔG/TT	175	44.3	39	24.1	1.4	0.45	
	TT/TT	176	44.6	116	71.6	4.1	0.0008	
	Total	395		162				
African-Americans, n=459								
rs12979860	TT	119	34.0	26	23.9	Ref.		0.58
	CT	183	52.3	54	49.5	1.35	0.26	
	CC	48	13.7	29	26.6	2.77	0.001	
	Total	350		109				
ss469415590	ΔG/ΔG	137	39.1	26	23.9	Ref.		0.62
	ΔG/TT	168	48	53	48.6	1.66	0.06	
	TT/TT	45	12.9	30	27.5	3.51	7.9E-05	
	Total	350		109				

OR = odds ratio for spontaneous clearance;
AUC = area under the receiver operating characteristic curve.

Supplementary Table 10. Association of the GWAS marker rs12979860 and ss469415590 with spontaneous clearance of hepatitis C virus infection among African-American injection drug users enrolled in ALIVE study. Analysis is limited to subjects successfully genotyped for both ss469415590 and rs12979860.

Variant	Genotype	Chronic	%	Clear	%	OR	p-value	AUC p-value
rs12979860	TT	228	38.9	15	19.2	Ref.		0.64
	CT	275	46.9	37	47.4	2.05	0.025	
	CC	83	14.2	26	33.3	4.76	7.61E-06	
	Total	586		78				
ss469415590	ΔG/ΔG	255	43.5	19	24.4	Ref.		0.64
	ΔG/TT	265	45.2	36	46.2	1.82	0.043	
	TT/TT	66	11.3	23	29.5	4.68	5.48E-06	
	Total	586		78				

OR = odds ratio for spontaneous clearance;
AUC = area under the receiver operating characteristic curve.

Supplementary Table 11. Primers and assays

primer	sequence	comments
IFNAN_cloning_F	ATGCGGCCGAGTGTCTGGGCC	additional overhangs were added for cloning
IFNAN_cloning_R	GAGGCAAGGCCCAGAGTGTGCAG	into specific vectors
IFNAN_F_seq_intr1_3'UTR	GTAAGTCACCGGTCAGCCCCTGTGCC	1397-bp amplicon for genotyping by sequencing of IFNAN
IFNAN_R_seq_intr1_3'UTR	CCCATTGACTGAGAGCCTCGCCCGG	variants within intron 1-exon 5
IFNAN_ex1_seq_F	CGAACCAGGGTTGAATTGC	628-bp amplicon for genotyping by sequencing of IFNAN
IFNAN_ex1_seq_R	GCACTGCAGACAGGAGTGAG	variants within exon 1 and 5'UTR
TaqMan expression assays with MGB probes		
ss469415590_IFNAN_F	GCCTGCTGCAGAAGCAGAGAT	ABI expression buffer
ss469415590_IFNAN_R	GCTCCAGCGAGCGGTAGTG	RNA must be DNAseI-treated
ss469415590_IFNAN_VIC (TT allele)	ATCGCAG**AA**GGCC	
ss469415590_IFNAN_FAM (ΔG, allele)	ATCGCAG**C**GGCCC	
IL28B_F	CGGAAGAGGTTGAAGGTGAC	ABI expression buffer
IL28B_R	CTCCACCATTGGCTGCAC	RNA must be DNAseI-treated
IL28B_probe (FAM)	GCCCCAAAAAAGGA	182 bp on DNA, 89 bp on cDNA template
IL29	assay Hs00601677_g1, Applied Biosystems	
PPIA (endogenous control)	assay 4326316E, Applied Biosystems	
TaqMan genotyping assays with MGB probes		
ss469415590_IFNAN_F	GCCTGCTGCAGAAGCAGAGAT	standard, Qiagen genotyping buffer
ss469415590_IFNAN_R	GCTCCAGCGAGCGGTAGTG	
ss469415590_IFNAN_VIC (TT, non-risk allele)	ATCGCAG**AA**GGCC	
ss469415590_IFNAN_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	