.

Multi-ancestry Fine Mapping of Interferon Lambda and the Outcome of Acute Hepatitis C Virus Infection

Candelaria Vergara[1], Priya Duggal[1], Chloe L. Thio[2], Ana Valencia[2,3], Thomas R. O'Brien[4], Rachel Latanich[2], Winston Timp[2], Eric O. Johnson[5], Alex H. Kral[5], Alessandra Mangia[6], James J. Goedert[7], Valeria Piazzola[6], Shruti H. Mehta[1], Gregory D. Kirk[1], Marion G. Peters[8], Sharyne M. Donfield[9], Brian R. Edlin[10], Michael P. Busch[11], Graeme Alexander[12], Edward L. Murphy[11], Arthur Y. Kim[13], Georg M. Lauer[14], Raymond T. Chung[14], Matthew E. Cramp[15], Andrea L. Cox[2], Salim I. Khakoo[16], Hugo R. Rosen[17], Laurent Alric[18], Sarah J. Wheelan[1,2], Genevieve L. Wojcik[19], David L. Thomas[2] and Margaret A. Taub[1].

[1] Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD. USA.

[2] Johns Hopkins University, School of Medicine, Baltimore, MD, USA

[3] Universidad Pontificia Bolivariana, Medellin, Colombia.

[4] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

[5] RTI International, Research Triangle Park, NC. USA

[6] Liver Unit IRCCS "Casa Sollievo della Sofferenza", San Giovanni Rotondo, Italy

[7] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD. USA

[8] Department of Gastroenterology, University of California, San Francisco, CA., USA.

[9] Rho, Inc., Chapel Hill, NC, USA

[10] Centers for Disease Control and Prevention, Atlanta, GA., USA

[11] University of California San Francisco and Vitalant Research Institute, San Francisco, CA., USA

.

[12] University College London Institute for Liver and Digestive Health, The Royal Free Hospital, London, UK.

[13] Division of Infectious Diseases, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA. USA

[14] Liver Center and Gastrointestinal Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA.,USA

[15] South West Liver Unit, Plymouth, PL6 8DH, UK.

[16] University of Southampton, Southampton General Hospital, Southampton, UK.

[17] University of Colorado, Aurora, CO., USA

[18] Department of Internal Medicine and Digestive Diseases, Centre Hospitalier Universitaire Rangueil, UMR 152, Institut de Recherche pour le Développement Toulouse 3 University, France.

[19] Department of Genetics, Stanford University School of Medicine, Stanford, CA., USA

Short title

Fine-mapping of *IFNL* Locus for HCV clearance

Corresponding Author

Candelaria Vergara, MD, PhD

725 N. Wolfe Street, Suite 211

Baltimore, Maryland 21205

Email: cvergar2@jhmi.edu

Disclosures: All authors declare that they have no conflict of interest and nothing to disclose.

.

Funding: R01 013324, R01-HD-41224, RO1-AI148049-21, R21AI139012, U19-AI082630, U01-AI131314, R01-DA033541, U19-AI066345, DA12568, DA036297, R01-DA09532, R01-DA11860, N02CP91027, H79TI12103, R01DA16159, R01DA21550, R01-HD-41224.

Abstract:

Clearance of acute infection with hepatitis C virus (HCV) is associated with the chr19q13.13 region containing the rs368234815 (TT/ΔG) polymorphism. We fine-mapped this region to detect possible causal variants that may contribute to HCV-clearance. First, we performed sequencing of *IFNL1-IFNL4* region in 64 individuals sampled according to rs368234815 genotype: TT/clearance (N=16) and ΔG/persistent (N=15) (genotype-outcome concordant) or TT/persistent (N=19) and ΔG/clearance (N=14) (discordant). 25 SNPs had a difference in counts of alternative allele > 5 between clearance and persistence individuals. Then, we evaluated those markers in an association analysis of HCV clearance conditioning on rs368234815 in two groups of European (692 clearance/1 025 persistence) and African ancestry (320 clearance/1 515 persistence) individuals. 10/25 variants were associated ($P < 0.05$) in the conditioned analysis leaded by rs4803221 ($P=4.9x10^{-04}$) and rs8099917 ($P=5.5x10^{-04}$). In the European ancestry group, individuals with the haplotype rs368234815ΔG/rs4803221C were 1.7x more likely to clear than those with the rs368234815ΔG/rs4803221G haplotype ($P=3.6x10^{-5}$). For another nearby SNP, the haplotype of rs368234815ΔG/rs8099917T was associated with HCV-clearance compared to rs368234815ΔG/rs8099917G (OR: 1.6, $P=1.8x10^{-4}$). We identified four possible causal variants: rs368234815, rs12982533, rs10612351 and rs4803221. Our results suggest a main signal of association represented by rs368234815, with contributions from rs4803221, and/or nearby SNPs including rs8099917.

Introduction

.

The outcome of the acute hepatitis C virus (HCV) infection is determined in part by host genetic factors. Previous genome-wide association studies (GWAS) and meta-analyses have identified significant associations of spontaneous clearance of HCV infection with several single nucleotide polymorphisms (SNPs) in the region harboring 4 interferon-λ genes (*IFNL1*, *IFNL2, IFNL3* and *IFNL4)* on chromosome 19q13.13 (1-3). Of particular importance is a dinucleotide variant in the first exon of *IFNL4*, rs368234815 (ΔG/TT), which causes a shift in the open reading frame of the gene; the presence of the ΔG allele at the variant position allows the expression of a fully functional IFNλ4 protein of 179 amino-acids (4,5). This allele is implicated in reduced HCV clearance (4). On the contrary, the TT allele is predicted to induce nonsense-mediated mRNA decay and is associated with increased HCV clearance (4) (Figure 1, Top panel).

Despite the strong and replicated association, some HCV infected individuals carrying the favorable genotype (TT) of rs368234815 do not clear the infection, while some patients with the unfavorable genotypes spontaneously clear the infection (4,6-10). This discordant *IFNL* genotype with HCV infection outcomes are not explained by other determinants of spontaneous clearance such as polymorphisms in other known HCV related genes, sex, or HIV co-infection. Thus, we reasoned that other variants in the *IFNL* region may contribute to the observed spontaneous clearance.

The *IFNL1-IFNL4* region is under strong linkage disequilibrium (LD) which complicates the identification of causal alleles (11). Moreover, sequencing of the region is limited by the presence of genes with high homology that precludes the accurate assignment of reads to a specific location (12,13). In this study we sought to overcome these challenges and to identify

.

variants that may contribute to clearance of HCV infection. We implemented two approaches to identify variants with an association independent of rs368234815 (Figure 2). First, we performed short-read sequencing analysis in a selected panel of individuals where the rs368234815 genotype was either concordant or discordant with the expected HCV outcome (clearance or persistence) using a sequencing strategy that allowed the precise assignment of the reads to specific coordinates in the locus and accurate calling of the variants. Second, we performed conditional analysis in a large independent set of individuals of European and African ancestry evaluated for display of HCV spontaneous clearance for whom we had genotypes imputed to the 1000 Genomes Project (14). To identify potential causal variants in the region we used well established statistical methods combining functional data from external sources with the association and LD patterns from our datasets (Figure 2). We also interrogated our dataset for association of two variants (rs1176648444 and rs4803217) that have been identified as functionally relevant in the region.

Materials and Methods

Genetic structure of the *IFNL* region: The interferon lambda region spans 50Kb of human chromosome 19q (15,16), Figure 3. The 4 interferon-λ genes seem to originate from gene duplication events (4,17) with *IFNL2* and *IFNL3* more closely related to each other than *IFNL1* (4). *IFNL1* and *IFNL2* are transcribed from the positive strand with a coding region of 2,348 and 1,445 basepairs (bp), with 5 and 6 exons, respectively. *IFNL3* and *IFNL4* have 6 and 5 exons each, are transcribed from the negative strand and have a coding region of 1 336 and 2 543 bp, respectively (Figure 3) (16).

Short-read sequencing in panel of discordant and concordant individuals

.

Individuals for IFNL Sequencing: Individuals included in this approach are part of the HCV Extended Genetics Consortium (2,18). Each individual study obtained consent for genetic testing from their governing Institutional Review Board and the overall research was approved by the Johns Hopkins School of Medicine Institutional Review Board (3). For this analysis we selected 64 individuals based on the genotype of rs368234815 and HCV spontaneous clearance/persistence status in a similar approach presented by Rauch et al for rs8099917 (1). This included 45 individuals of African Ancestry (21 clearance/24 persistence) and 19 individuals of European Ancestry (9 clearance/10 persistence) (Table 1, Figure 1, Bottom panel). These individuals were either concordant between genotype and HCV outcome (i.e. favorable genotype of rs368234815 [TT/TT] and HCV spontaneous clearance or unfavorable genotypes rs368234815 [ΔG/ΔG] and HCV persistence) or were discordant (i.e. favorable genotype and HCV persistence or unfavorable genotype and HCV clearance) (Table 1, Figure 1, Bottom panel).

IFNL Sequencing: Because the region containing the four *IFNL* genes has low sequence complexity, the alignment of short reads generated with standard high-throughput sequencing methods is challenging (12); thus, we designed a targeted sequencing approach where the entire 70.8 Kb region (chromosome 19:39721399-39792284, coordinates based on The Genome Reference Consortium Human build 37-GRCh37-) was amplified in eight segments with customized primers (Figure 3, Supplementary Table 1). This allowed alignment of reads specifically to the region of origin, resulting in more confident detection of individual variants across the whole region. Methods of DNA extraction and strategies for sequencing and alignment are described in detail in Supplementary Material.

.

Statistical Analysis of the sequenced panel: Counts of alternative (non-reference, non-ancestral) alleles at each position of the sequenced region were generated to compare differences in single-nucleotide variants (SNVs) between concordant and discordant individuals. We report all positions where the difference in alternative allele count is at least 5. Given sample size limitations, we did not perform formal statistical tests for differences between groups, but report all positions for validation in imputation data. Customized scripts in R (https://www.r-project.org/) were used to do the SNV analysis. Results of comparison of all variants included in the analysis of the region is available upon request.

Conditional Analysis of an independent imputed dataset.

Individuals of the independent imputed dataset: We analyzed the rest of the individuals of the HCV Extended Genetics Consortium (2,18), corresponding to an independent set of 1835 individuals of African Ancestry (320 clearance/1515 persistence) and 1717 individuals of European ancestry (692 clearance/1025 persistence). The rs368234815 ΔG allele had a frequency of 0.63 in the African ancestry group and 0.31 in the European ancestry group (Table 1).

Genotyping and Imputation: Genotypes in this region were derived from a genome-wide association study previously described (2,3) and in Supplementary Material. For this analysis, 421 high quality imputed variants in the *IFNL* region were used in African ancestry individuals and 282 in European ancestry individuals.

Statistical Analysis of the independent imputed dataset: In each of the ancestry groups, we performed an association analysis of dosage of the variants in the region conditioned on the rs368234815 variant using an additive logistic regression model, adjusting for 3 principal

.

components and HIV status using Mach2dat (19). Conditional analysis in the two ancestry groups were meta-analyzed using the fixed effects inverse variance method in METAL software (20). Given that this region has been highly replicated and to preserve power for detect secondary signals in the fine mapping, a value of $P < 0.05$ was considered as significant (21). Results from the imputation analysis were pulled to check for significance at any of the loci that were identified based on alternative allele counts from the sequencing analysis. Candidate sites with difference in the allele count and significance in the imputed dataset were carried forward for the haplotype analysis.

Haplotype analysis: To further characterize the locus across populations, we conducted haplotype association analyses. We calculated LD and constructed haplotypes based on the candidate sites in the European ancestry and African ancestry populations. LD patterns in each population were calculated using the algorithm from Gabriel et al (22) in Haploview (23) in the imputed dataset. We performed haplotype analyses using the "haplo.stat" R package (24). We assumed an additive model in which the regression coefficient represented the expected change in the log odds of HCV clearance with each additional copy of the specific haplotype compared with the reference haplotype.

Identification of potential causal variants

To identify variants that may be causal or have a regulatory function we refined the region observed in a previous GWAS of HCV clearance (3). We used association summary statistics from GWAS for those markers in the *IFNL* region, leveraged functional data and LD information of the included markers and described a 99% credible set of variants using PAINTOR (25,26) as described in detail in Supplementary Materials. Aiming to optimize power for this analysis, we included the complete dataset of the HCV Extended Genetic consortium

.

comprising 3608 people from two ancestry groups: 1869 individuals of African ancestry (340 clearance/1529 persistence) and 1736 of European ancestry (701 clearance/1,035 persistence). We considered PAINTOR predicted variants to be functional based on a posterior probability > 0.1, a threshold suggested previously (26). To investigate functional elements, the presence or absence of overlap was determined by the UCSC Table Browser intersecting the calculated credible set with the signal tracks described in Supplementary Materials.

Analysis of functionally relevant variants

Two markers (Rs4803217 and Rs1176648444) has been described as modulators of the association given by rs368234815 (Supplementary Material). Given their potential functional role, we evaluated their allele count in the sequenced dataset and the association of each variant in the imputed dataset after conditioning on rs368234815. We also evaluated the residual association after conditioning on both rs368234815 and rs1176648444 and the association of the rs368234815- rs4803217 and rs368234815- rs1176648444 haplotypes in each population and using the methods described in haplotype analysis, we constructed and evaluated association of haplotypes based on the candidate sites common to European and African ancestry populations incorporating these functionally relevant variants.

Results

When analyzing all individuals of the sequencing group (concordants and discordants), we identified 25 positions (candidate SNVs) where the difference in the frequency of the alternative allele was >5 between the individuals with clearance and persistence (Table 2). The identified variants are located downstream of *IFNL3-IFNL4* and in the intergenic regions between *IFNL4-IFNL2* and *IFNL2-IFNL1* (Supplementary Figure 1).

.

Conditional Analysis in an independent imputed dataset. From all the variants analyzed in the region in this dataset, we extracted the results of the 25 variants identified in the targeted sequencing panel. From those, 2 variants were not present in both ancestry groups and 10 candidate variants were significantly associated (P value < 0.05) in the meta-analysis (Supplementary Figure 1, Table 2). No other variants in the region was significantly associated in the conditional analysis.

The 10 candidate variants are located in an 11.3Kb region (chr19: 39732501-39743821) spanning 1.7 Kb downstream of *IFNL3* and 4.3Kb upstream of *IFNL4* (Figure 4, Panel A-B and Supplementary Figure 1). The association observed in this meta-analysis was driven mainly by the contribution of the European ancestry samples (Table 2). In this group, nine of ten associated variants have similar minor allele frequencies (~ 0.19) with the exception being rs4803222 (0.30). Rs4803221, a synonymous SNP in *IFNL4* (NM_001276254: S (TC**G**) --> S (TC**C**)), had the strongest association in the meta-analysis (P-value = 4.86×10^{-04}) and in the European ancestry group (P-value = 7.49×10^{-06}). In the African ancestry population, the direction of the effect is the same but no variants were significantly associated. The allele frequencies for seven of ten variants were similar across the ancestry groups. However, at rs8107030, rs8099917 and rs7248668 minor allele frequencies were lower in the African ancestry population (0.04-0.06) than in persons of European ancestry (0.19) (Table 2).

Haplotype analysis. The markers included in the haplotype analysis were rs368234815 and the ten candidate variants (Figure 4, Panel C). Haplotype construction revealed that in the European ancestry individuals all 11 SNPs are within one unique haplotype block of 11kb with LD values consistently high ($r^2 > 0.89$), except for rs4803222 and rs368234815 with an $r^2 \sim 0.50$ with the

.

other variants (Figure 4). Thus, the top SNP of the candidate variants (rs4803221) tags all associated variants in the block except for rs4803222 and rs368234815.

We estimated the frequency of the haplotypes based on the boundaries determined in the haplotype blocks for each population. In the European ancestry samples, the 11 markers formed 24 haplotypes of which four (denoted H1 to H4) had a frequency higher than 2 % so were included in the haplotype association analysis (Supplementary Table 2). H1 (containing the favorable TT allele of rs368234815) was the haplotype with highest prevalence overall and was more frequent in the clearance group (P-value= 4.4×10^{-22}). H2, H3 and H4 contained the unfavorable allele (ΔG) of rs368234815. H3 and H4 and were significantly associated with persistence (P value < 0.05. H2 had a low prevalence (2%) and was not associated with HCV clearance (P value = 0.53).

In African ancestry individuals, there was unique haplotype block of 5Kb containing 9 out of the 11 variants with LD r^2 values ranging from 0.99 to 0.03 (Figure 4). In this population rs4803221 is able to capture information only from rs8105790, rs66531907, rs12983038, rs8109889 but not from rs8099917 or rs7248668. Similarly, rs8107030, rs368234815 and rs4803222 only capture information from themselves. Similar to the European ancestry population, rs368234815 had low LD r^2 values with the ten variants (Figure 4).

In African ancestry individuals, 21 haplotypes were present and five with the highest frequency (> 6%) were included in the analysis. Haplotypes H1-H4 are similar to those of the European ancestry populations for the shared markers. Similarly, H1 was more frequent in the clearance group (P value= 1.5×10^{-14}), however, this haplotype had a considerably lower prevalence in this

.

246 sample compared to the European ancestry sample (37.5% vs. 68.1%), which can be explained
247 by the differences in the allelic frequency of the ΔG allele between those samples (Table 1).

248 In the African ancestry sample, H2 was the predominant haplotype conferring persistence and
249 has a considerably higher prevalence in this African ancestry group (36% vs 2% in the European
250 group). H3 and H4 were associated with persistence with comparable effect size but lower
251 significance (P = 0.02). In summary, in both sample groups H1 was significantly associated with
252 clearance. However, the predominant haplotypes conferring persistence were different in each
253 sample group (H3 in European ancestry vs. H2 in African ancestry individuals), Supplementary
254 Table 2. Similar results were observed in the haplotype analysis including common candidate
255 variants and functionally relevant variants, except for the separation of H2 in African Ancestry
256 individuals in H2a and H2b. H2b conserved similar direction and strength of effect than H2
257 (Supplementary Table 3) .

258 Next, to determine whether an allele or haplotype could "overcome" the unfavorable allele (ΔG)
259 of rs368234815, we restricted our analysis to haplotypes containing this unfavorable allele in the
260 European ancestry group. We found that the haplotypes with the C allele at rs4803221 were
261 significantly associated with clearance compared to those containing the G allele (OR: 1.7, 95%
262 CI: 1.3-2.29, P value = 3.6×10^{-5}, Table 3). This was not observed in African ancestry
263 individuals (OR for haplotype with the C allele: 1.25; 95% CI: 0.81-1.9, P value = 0.29).
264 Rs4803221 tags rs8099917 and rs7248668 in the European ancestry group but not in the African
265 ancestry group. Similar to rs4803221, the haplotype containing the T allele of rs8099917 (and G
266 allele of rs7248668) is significantly associated with clearance compared to the one containing the
267 G allele of rs8099917 (and A allele of rs7248668: OR: 1.6, 95% CI: 1.3-2.16, P value: 1.76x 10-
268 4) in the European ancestry individuals but not in African ancestry individuals (Table 3). These

.

data are consistent with a main signal being shared across populations driven by one or more functional variants represented by rs368234815, with potential additional contributions from rs4803221, and/or proxies including rs8099917 and rs7248668 in the European ancestry population.

Identification of potential causal variants

Four SNPs were identified as likely functional (posterior probability > 0.1, Supplementary Table 4). The credible set obtained with PAINTOR, determined by 2 out of 4 variants of the credible set (rs368234815 and rs12982533), overlaps with the previously estimated credible set using a larger dataset (3), narrowing the signal to a 7251 bp region (19:39731904-39739155) located 2368bp downstream from *IFNL3* and extending until exon 1 of *IFNL4* (Supplementary Figure 2). The identified region includes rs368234815 which we confirmed as the main driver of the association signal. The variants identified in the fine-mapping credible set overlapped with regulatory regions in hepatocyte cell lines and liver tissue including CpG sites that are completely or partially methylated, target sites for transcription factors, DNA methylation sites with 50-100% methylation in those cells, candidate weak enhancers, polycomb repressors and with transcription associated activity (Supplementary Figure 2). Two other variants (rs4803221 and rs10612351) also showed posterior probability values > 0.10 indicating that they might be considered causal even though rs10612351 is not included in calculated 99% credible set. We considered that these polymorphisms are plausible candidate variants based both on fine-mapping and regulatory overlap and these results support the findings of the haplotype analysis.

Analysis of functionally relevant variants

.

Rs4803217 and rs1176648444 had a difference of 1 and 0 respectively in the counts of the alternative allele between clearance and persistence in the sequenced panel. Rs4803217 showed no association in the imputed dataset after conditioned on rs368234815 (European ancestry conditioned P value= 0.15, African ancestry conditioned P value= 0.3, Meta-analysis conditioned P value= 0.08). Rs1176648444 was not associated in African Ancestry (P value= 0.38) but interestingly it showed a significant association in individuals of European Ancestry only in the conditioned analysis (Not conditioned P value= 0.07; conditioned P value=0.00003, conditioned meta-analysis P value = 0.06). In the double conditioned analysis with rs368234815 and rs1176648444, six out of ten variants associated in the single conditioned analysis showed residual association (Supplementary Table 5). Rs368234815TT- Rs4803217C haplotype was significantly associated with clearance compared with the Rs368234815ΔG-Rs4803217A in both populations. In African ancestry population the haplotype Rs368234815ΔG-Rs4803217C was significantly associated with persistence (Supplementary Table 6). On the other hand, in the European ancestry population, the haplotype rs368234815ΔG- rs1176648444A (IFNλ4-S70) was associated with clearance with an intermediate effect between rs368234815ΔG-rs1176648444G (IFNλ4-P70) and rs368234815TT-rs1176648444G (no IFNλ4), Supplementary Table 7.

Discussion

We performed a comprehensive, trans-ethnic analysis of genetic variation in the *IFNL* region and spontaneous recovery from HCV infection. We discovered variants with associations independent of the well-described rs368234815 variant that suggest additional genetic contributions to the outcome of this chronic infection.

We observed an rs368234815-independent signal led by rs4803221 (given mainly for the European Ancestry population) and ten other variants in LD including rs8099917 and rs7248668.

.

In sensitivity analysis we confirmed that this signal was present even after conditioning on rs368234815 and rs1176648444 indicating a residual or modifying effect of the remaining variants**.** The LD structure of the region in the European ancestry group suggests that the rs4803221 association may be due to any one of a number of variants including rs8099917 and rs7248668. In fact, in the context of haplotypes conferring persistence in this group, the haplotype containing the C allele of rs4803221, the T allele of rs8099917 and the G allele of rs7248668 were significantly associated with HCV clearance. However, we do not observe a significant signal in individuals of African ancestry at rs4803221, even though its allele frequency and the sample size are similar to the European ancestry group. It is possible that the association of rs4803221 observed in the European ancestry group is explained by linkage with rs8099917 and/or rs7248668, instead of being functional itself. Unfortunately, our power was limited to confirm this inference in the African ancestry group, where we detected an odds ratio of 0.79 with a MAF of 0.06 (power of only 0.38 at a significance level of 0.05, compared to 1 in the European ancestry group) (27).

Rs4803221 is a variant with multiple functions which has been previously linked to HCV spontaneous clearance in individuals with beta thalassemia (28). The SNP is located in exon one of *IFNL4*, 357 bp downstream from the transcription start site and 3522 bp upstream from the transcription start site of *IFNL3*; the G allele (MAF=0.2) abolishes a CpG site and induces a synonymous (Ser>Ser) change at position 30 of the *IFNL4* protein. Similar to our findings, Origa *et al,* found an association with rs4803221 that was independent of rs12979860 (which is itself in high LD with rs368234815) (28) . Rs4803221 significantly improved the viral clearance prediction in patients carrying the un-favorable T allele of rs12979860 (in high LD with the un-favorable ΔG allele of rs368234815). They hypothesized that the abolishment of methylation

.

sites might increase expression of *IFNL3* and downregulate interferon sensitive genes, reducing net innate antiviral activity (28).

A potential 'causal' role has also been described for rs8099917 (1). In a GWAS including 1362 European ancestry individuals, G allele was associated with persistence of HCV infection (1). Several specific SNPs were identified as candidates for being causal, however rs368234815 was not described in this panel. In European HCV-infected individuals analyzed for response treatment, haplotypes tagged by the T allele of rs8099917 showed higher expression of IFNλ3 and IFNλ2 but no evaluation was reported on expression of IFNλ4 (29). Analogous results were found in a Japanese cohort where the expression *IFNL3* and *IFNL2* mRNA was lower in the carriers of the G allele (30). In our current study, T allele is associated with clearance in the context of the ΔG allele of rs368234815, which corresponds with *IFNL4* transcription but HCV persistence (4). Rs8099917 is located 8.9 kb upstream from *IFNL3* and 16 kb upstream from *IFNL2*. If we assume the model that rs368234815 regulates the expression of IFNλ4 (Figure 1), it would be worthwhile to investigate if the statistically independent effect of rs8099917 observed in this study is perhaps caused by an increase in expression of IFNλ3 and IFNλ2, a decrease in the production of IFNλ4 in those individuals with the ΔG allele at rs368234815, or both.

SNP rs7248668 located in the 5' region of *IFNL4* is in high LD with rs8099917 in populations included in this analysis and in The 1000 Genomes Project independently of ancestry (14). In fine mapping analysis, the haplotype containing the G allele was associated with virologic response to pegylated interferon-α and ribavirin therapy for chronic hepatitis C in a Japanese population (30). Similarly, the patients with the GG genotype showed virologic response rates up to four times higher than those for patients with unfavorable genotypes in HIV/HCV co-infected

.

patients of European ancestry (31). In our study the same G allele is associated with spontaneous clearance; even though the phenotypes are not completely comparable, in general the G allele favors the clearance of the virus in each context across studies. Due to their high LD the effect of rs7248668 is not separable from that of rs8099917.

Our fine mapping analysis using PAINTOR indicates that the potential causal variant in this locus is contained in the *IFNL3-IFNL4* gene region. This credible set informed by our analysis harbors the compound di-nucleotide exonic variant (rs368234815, ΔG/TT) and the rs4803221 variant, but does not contain rs8099917 or rs7248668. It is possible that we did not find a high posterior probability for those 2 latter SNPs because they are in high LD with rs4803221 in European ancestry subjects, where the significant independent effect was observed. We consider that the expansion of the sample size of African ancestry individuals could allow disentanglement of the effects of rs4803221, rs8099917 and rs7248668. The coding nature of the rs368234815, the high significance and large effect-size, and the low LD between this variant and the others in the region (especially in the African ancestry population) contributed to determine this variant as functionally relevant.

The identification of rs12982533 as functionally relevant deserves further analysis. This variant has been included as part of haplotypes associated with response to treatment (13,32) but not with spontaneous HCV clearance; it is located 3.7kb 3' of *IFNL3* and its functional role is unknown. It is important to notice that we limited this analysis to only variants that were consistently present in both ancestry groups and it is possible that this set of variants fails to capture putatively important variation within or around the *IFNL* locus.

.

The results of the of rs368234815-rs1176648444 haplotype analysis in European ancestry agree with findings previously described by the Swiss Hepatitis Cohort Study Group (33) where they demonstrated that individuals with IFNλ4-S70 have rates of HCV clearance that are intermediate to those with IFNλ4-P70 and those with rs368234815TT/TT genotype, who do not produce the IFNλ4 protein. Similarly, our findings on rs368234815-rs4803217 haplotype are in concordance with the association of rs368234815ΔG: rs4803217G with the poorest virologic response to peg– interferon alpha and ribavirin therapy in African Americans (34); even though it is not the same phenotype, it suggests an interaction of the two variants responsible for lower rate of resolution of the infection in general. The restriction of the haplotype effect to specific populations deserves further analysis including a larger sample capable of capturing all haplotype diversity.

One strength of this analysis is the sequencing strategy which allowed us to unambiguously map read pairs to specific segments of the *IFNL* region, and call the genetic variants with higher accuracy than using conventional methods of short-read sequencing. The "conditioned by design" composition of the panel with concordant and discordant individuals enabled the detection of variants conferring an effect on HCV clearance that is adjusted for the allele present at rs368234815. Even though the sample size of the sequencing panel is small, its particular configuration makes it suitable to detect variants with a large effect. The findings from this panel were supported by the results of the statistically conditioned analysis in a much larger sample size with similar characteristics adding reliability to the findings. One limitation of the study is that we established a rather high cut off for the selection of the variants with differences in the allele count since the size of the panel precluded the evaluation of rare variants using standard

.

statistical tests and we excluded rare variants in the imputation panel since the imputation quality is usually low for those variants and any derived results would be considered uncertain.

In this study we fine-mapped the *IFNL* region and found results that support an independent genetic effect of several variants in this locus. Our results are applicable to the European ancestry population with our current sample sizes and are hypothesis-generating regarding additional factors contributing to the higher clearance in European ancestry and African ancestry individuals. Our findings are relevant and complementary to previous analyses aimed to understand the genetic basis of HCV clearance and the differences in the immune response to this infection across populations.

References

(1) Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 2010 Apr;138(4):1338-45, 1345.e1-7.

(2) Duggal P, Thio CL, Wojcik GL, Goedert JJ, Mangia A, Latanich R, et al. Genome-wide association study of spontaneous resolution of hepatitis C virus infection: data from multiple cohorts. Ann Intern Med 2013 Feb 19;158(4):235-245.

(3) Vergara C, Thio CL, Johnson E, Kral AH, O'Brien TR, Goedert JJ, et al. Multi-Ancestry Genome-Wide Association Study of Spontaneous Clearance of Hepatitis C Virus. Gastroenterology 2018 Dec 26.

.

(4) Prokunina-Olsson L, Muchmore B, Tang W, Pfeiffer RM, Park H, Dickensheets H, et al. A variant upstream of IFNL3 (IL28B) creating a new interferon gene *IFNL4* is associated with impaired clearance of hepatitis C virus. Nat Genet 2013 Feb;45(2):164-171.

(5) Hong M, Schwerk J, Lim C, Kell A, Jarret A, Pangallo J, et al. Interferon lambda 4 expression is suppressed by the host during viral infection. J Exp Med 2016 Nov 14;213(12):2539-2552.

(6) Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009 Oct 8;461(7265):798-801.

(7) Tillmann HL, Thompson AJ, Patel K, Wiese M, Tenckhoff H, Nischalke HD, et al. A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice. Gastroenterology 2010 Nov;139(5):1586-92, 1592.e1.

(8) Knapp S, Warshow U, Ho KM, Hegazy D, Little AM, Fowell A, et al. A polymorphism in IL28B distinguishes exposed, uninfected individuals from spontaneous resolvers of HCV infection. Gastroenterology 2011 Jul;141(1):320-5, 325.e1-2.

(9) Aka PV, Kuniholm MH, Pfeiffer RM, Wang AS, Tang W, Chen S, et al. Association of the IFNL4-DeltaG Allele With Impaired Spontaneous Clearance of Hepatitis C Virus. J Infect Dis 2014 Feb 1;209(3):350-354.

(10) Chinnaswamy S. Genetic variants at the IFNL3 locus and their association with hepatitis C virus infections reveal novel insights into host-virus interactions. J Interferon Cytokine Res 2014 Jul;34(7):479-497.

.

(11) Spain SL, Barrett JC. Strategies for fine-mapping complex traits. Hum Mol Genet 2015 Oct 15;24(R1):R111-9.

(12) Mandelker D, Schmidt RJ, Ankala A, McDonald Gibson K, Bowser M, Sharma H, et al. Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing. Genet Med 2016 Dec;18(12):1282-1289.

(13) Smith KR, Suppiah V, O'Connor K, Berg T, Weltman M, Abate ML, et al. Identification of improved IL28B SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort. Genome Med 2011 Aug 31;3(8):57.

(14) Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. Nature 2015 Oct 1;526(7571):75-81.

(15) Kotenko SV, Gallagher G, Baurin VV, Lewis-Antes A, Shen M, Shah NK, et al. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. Nat Immunol 2003 Jan;4(1):69-77.

(16) Sheppard P, Kindsvogel W, Xu W, Henderson K, Schlutsmeyer S, Whitmore TE, et al. IL-28, IL-29 and their class II cytokine receptor IL-28R. Nat Immunol 2003 Jan;4(1):63-68.

(17) O'Brien TR, Prokunina-Olsson L, Donnelly RP. IFN-lambda4: the paradoxical new member of the interferon lambda family. J Interferon Cytokine Res 2014 Nov;34(11):829-838.

.

(18) Wojcik GL, Thio CL, Kao WH, Latanich R, Goedert JJ, Mehta SH, et al. Admixture analysis of spontaneous hepatitis C virus clearance in individuals of African descent. Genes Immun 2014 Apr;15(4):241-246.

(19) Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. Annu Rev Genomics Hum Genet 2009;10:387-406.

(20) Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010 Sep 1;26(17):2190-2191.

(21) Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet 2018 Aug;19(8):491-504.

(22) Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002 Jun 21;296(5576):2225-2229.

(23) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005 Jan 15;21(2):263-265.

(24) Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002 Feb;70(2):425-434.

(25) Kichaev G, Pasaniuc B. Leveraging Functional-Annotation Data in Trans-ethnic Fine-Mapping Studies. Am J Hum Genet 2015 Aug 6;97(2):260-271.

.

(26) Kichaev G, Yang WY, Lindstrom S, Hormozdiari F, Eskin E, Price AL, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet 2014 Oct 30;10(10):e1004722.

(27) Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003 Jan;19(1):149-150.

(28) Origa R, Marceddu G, Danjou F, Perseu L, Satta S, Demartis FR, et al. IFNL3 polymorphisms and HCV infection in patients with beta thalassemia. Ann Hepatol 2015 May-Jun;14(3):389-395.

(29) Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 2009 Oct;41(10):1100-1104.

(30) Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009 Oct;41(10):1105-1109.

(31) Fernandez-Rodriguez A, Rallon N, Berenguer J, Jimenez-Sousa MA, Cosin J, Guzman-Fulgencio M, et al. Analysis of IL28B alleles with virologic response patterns and plasma cytokine levels in HIV/HCV-coinfected patients. AIDS 2013 Jan 14;27(2):163-173.

(32) Booth DR, Ahlenstiel G, George J. Pharmacogenomics of hepatitis C infections: personalizing therapy. Genome Med 2012 Dec 26;4(12):99.

.

(33) Terczynska-Dyla E, Bibert S, Duong FH, Krol I, Jorgensen S, Collinet E, et al. Reduced IFNlambda4 activity is associated with improved HCV clearance and reduced expression of interferon-stimulated genes. Nat Commun 2014 Dec 23;5:5699.

(34) O'Brien TR, Pfeiffer RM, Paquin A, Lang Kuhs KA, Chen S, Bonkovsky HL, et al. Comparison of functional variants in IFNL4 and λL3 for association with HCV clearance. J Hepatol 2015 Nov;63(5):1103-1110.

Acknowledgments: We thank the participants of the study. We also thank Cristian Velarde for graphic design of the Figures.

Legends of Tables

Table 1. Genetic ancestry, HCV status and rs368234815 genotype distribution of the analyzed individuals.

Table 2. Variants with a difference ≥ 5 in alternative allele count in sequenced individuals and replication in the meta-analysis of the association test of imputed variants in the *IFNL* region conditioned on the rs368234815 genotype. Bold text indicates positions with meta-analysis $p<0.05$ from imputed data.

Table 3. Association analysis of the 2 variant haplotypes (rs4803221-rs368234815 and rs368234815- rs8099917) in individuals carrying the ΔG allele of rs368234815.

Legends of Figures

.

Figure 1. Top panel: Depiction of putative role of rs368234815 genotypes in HCV persistence or clearance. Bottom panel: Schematic representation of concordant and discordant panels of individuals used in the sequencing analysis.

Figure 2. Schematic representation of the fine mapping analysis performed in this study.

Figure 3. Genetic structure of the *IFNL* locus, amplified fragments used for targeted sequencing, and two main variants associated in prior GWAS studies with HCV clearance (rs12979860 and rs368234815). Genetic coordinates are based on The Genome Reference Consortium Human build 37 (GRCh37/hg19).

Figure 4. Results of the analysis conditioned on rs368234815 and LD patterns of the top associated variants from that analysis. A) Meta-analysis conditioned on rs368234815 genotype. Variants represented in squares in panel B are the 23 of 25 variants that showed a difference of at least 5 in counts between clearance and persistence groups in the sequencing analysis. Recombination in this region is plotted in the background in light blue. Pairwise LD between the top associated variant and other variants in the region were estimated using LD data in the European (EUR) population in the 1000 Genomes project (hg19/Nov 2014). The color from blue to red represents the r^2 values relative to the peak position after conditioning, rs4803221. B) P values of rs368234815 and the 10 SNPs with remaining significance in the conditional analysis and their location on the genes in the region. C) LD plot of those variants in individuals of European and African ancestry in the genotyped/imputed dataset. The value within each diamond of the LD plot represents the pairwise correlation between tagging SNPs defined by sides of each the diamond. Shading represents the magnitude and significance of pairwise LD represented by the r^2 value, with a red- yellow gradient reflecting higher to lower LD values.

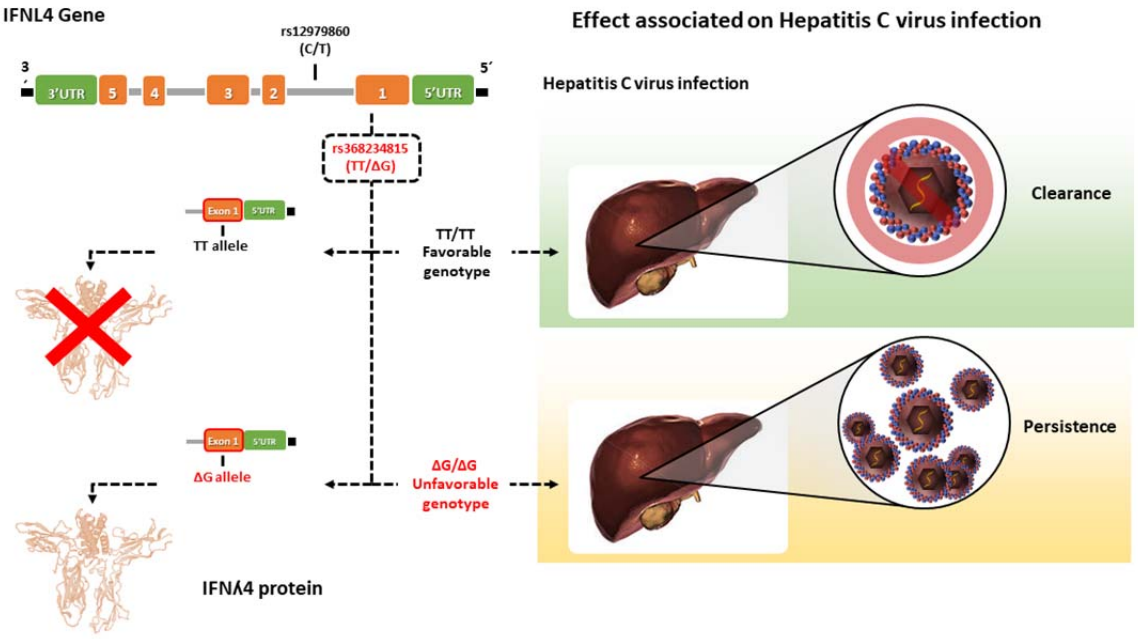
.

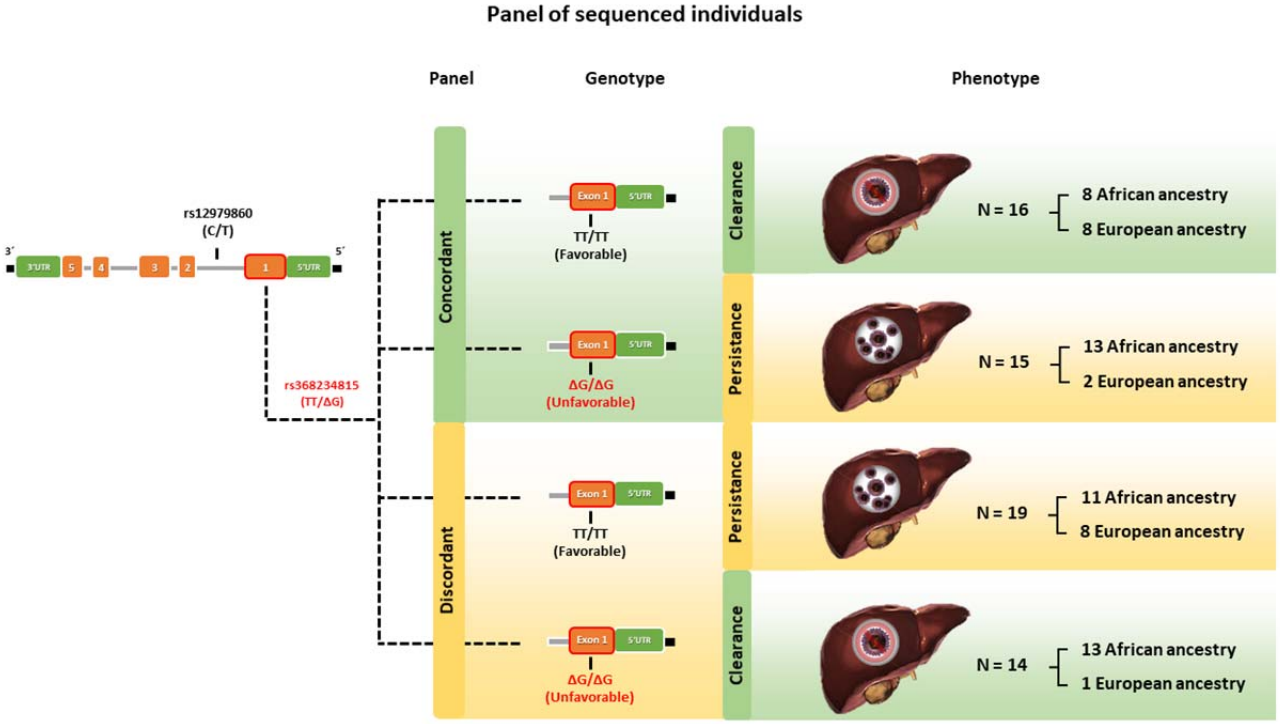
Association plots were graphed with Locus Zoom, P value and LD plots were generated using the package snp.plotter implemented in R.

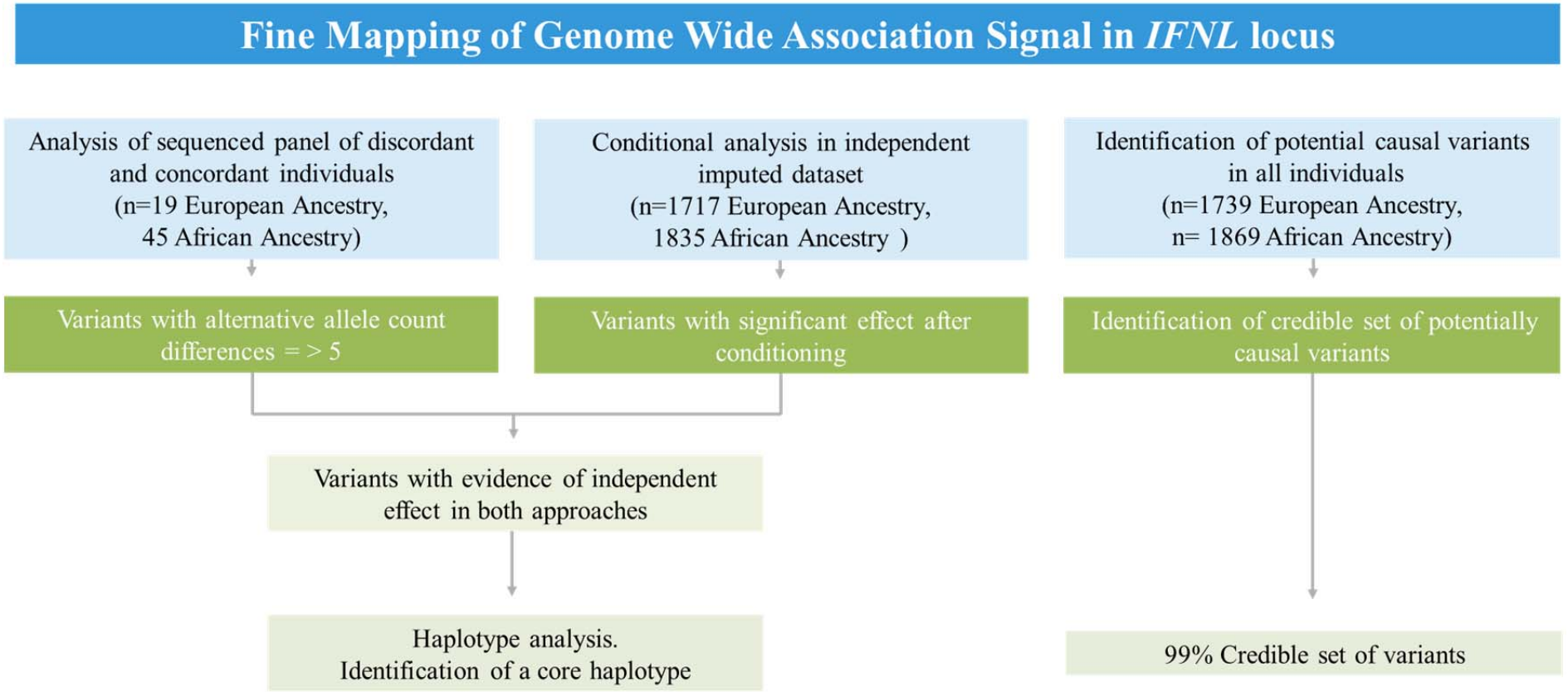
Ancestry group	HCV status	Sequencing analysis (N=64)			Imputation analysis (N=3552)			
		Genotype at rs368234815 (n)		Total	Genotype at rs368234815 (n)			Total
		ΔG/ΔG	TT/TT		ΔG/ΔG	TT/ΔG	TT/TT	
African Ancestry	Clearance	13	8	21	91	137	92	320
	Persistence	13	11	24	626	742	147	1515
	Total	26	19	45	717	879	239	1835
European Ancestry	Clearance	1	8	9	39	232	421	692
	Persistence	2	8	10	128	521	376	1025
	Total	3	16	19	167	753	797	1717

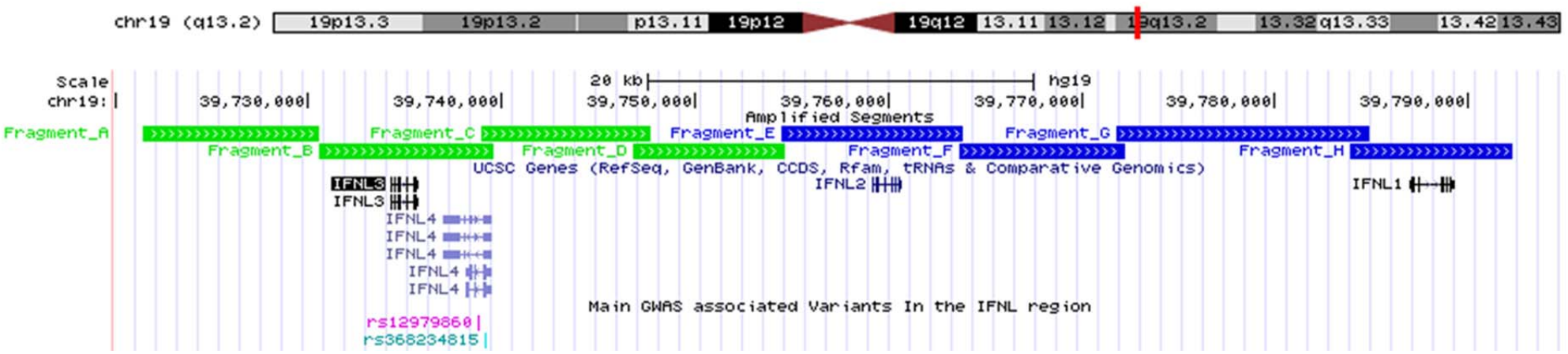
SNP				Analysis of individuals in sequencing panel (N=64)			Analysis of imputed data conditioned on rs368234815						
				Counts of Alternative allele			European Ancestry population (N=1,717)			African Ancestry Population (N=1,835)			Meta-analysis (N=3,552)
rsID	**Position**	**Ref**	**Alt**	**Clear**	**Persist**	**Diff.**	**Freq.**	**OR**	**P Value**	**Freq.**	**OR**	**P Value**	**P value**
rs8107090	39721915	T	A	19	24	-5	0.40	0.96	0.67	0.57	0.90	0.45	0.400
rs35408086	39726810	G	A	11	19	-8	0.39	0.98	0.80	0.21	0.95	0.66	0.621
rs11883239	39727480	G	A	6	17	-11	0.39	0.98	0.80	0.13	0.83	0.21	0.287
rs11883201	39727490	A	G	19	24	-5	0.40	0.97	0.69	0.59	0.93	0.59	0.507
rs955155	39729479	G	A	2	8	-6	0.26	0.98	0.84	0.07	0.86	0.41	0.459
rs12609937	39731204	A	G	28	35	-7	0.91	0.95	0.72	0.98	0.99	0.97	0.786
rs115166799	39732212	A	G	12	6	6	N/A	N/A	N/A	0.19	0.98	0.88	0.882
rs8105790	**39732501**	**T**	**C**	**6**	**13**	**-7**	**0.20**	**0.53**	**$3.21x10^{-05}$**	**0.19**	**0.94**	**0.64**	**0.001**
rs8102358	39735012	G	A	13	7	6	NA	N/A	NA	0.25	1.03	0.82	0.820
rs8107030	**39736719**	**A**	**G**	**0**	**7**	**-7**	**0.19**	**0.54**	**$4.84x10^{-05}$**	**0.04**	**0.91**	**0.72**	**0.002**
rs12971396	**39737866**	**C**	**G**	**7**	**13**	**-6**	**0.20**	**0.53**	**$2.79x10^{-05}$**	**0.19**	**0.94**	**0.68**	**0.001**
rs4803221	**39739129**	**C**	**G**	**7**	**13**	**-6**	**0.20**	**0.51**	**$7.49x10^{-06}$**	**0.19**	**0.93**	**0.60**	**$4.86x10^{-04}$**
rs73555604	39739170	C	T	12	6	6	0.01	1.31	0.47	0.22	1.01	0.97	0.597
rs4803222	**39739353**	**G**	**C**	**9**	**14**	**-5**	**0.30**	**0.55**	**0.06**	**0.27**	**0.85**	**0.23**	**0.029**
rs66531907	**39740675**	**C**	**A**	**6**	**12**	**-6**	**0.19**	**0.55**	**$5.90x10^{-05}$**	**0.19**	**0.90**	**0.48**	**$9.57x10^{-04}$**
rs12983038	**39741124**	**G**	**A**	**6**	**11**	**-5**	**0.19**	**0.54**	**$3.22x10^{-05}$**	**0.19**	**0.91**	**0.54**	**$8.75x10^{-04}$**
rs8109889	**39742770**	**C**	**T**	**6**	**12**	**-6**	**0.19**	**0.57**	**$1.01x10^{-04}$**	**0.19**	**0.90**	**0.46**	**0.001**
rs8099917	**39743165**	**T**	**G**	**0**	**5**	**-5**	**0.19**	**0.57**	**$1.22x10^{-04}$**	**0.06**	**0.79**	**0.28**	**$5.54x10^{-04}$**
rs7248668	**39743821**	**G**	**A**	**0**	**5**	**-5**	**0.19**	**0.57**	**$9.41x10^{-05}$**	**0.06**	**0.79**	**0.29**	**$5.14x10^{-04}$**
rs10853728	39745146	C	G	26	34	-8	0.65	0.94	0.58	0.74	0.85	0.18	0.177
rs10775535	39745181	C	T	29	34	-5	N/A	N/A	N/A	N/A	N/A	N/A	N/A
rs56116812	39747090	G	A	11	18	-7	0.12	1.34	0.06	0.23	0.90	0.41	0.473
rs116236518	39749790	C	T	5	0	5	N/A	N/A	N/A	0.02	1.48	0.26	0.262
rs10424607	39749922	A	C	18	23	-5	0.29	1.04	0.81	0.51	0.98	0.86	0.966
rs251908	39764449	A	G	30	35	-5	N/A	N/A	N/A	N/A	N/A	N/A	N/A

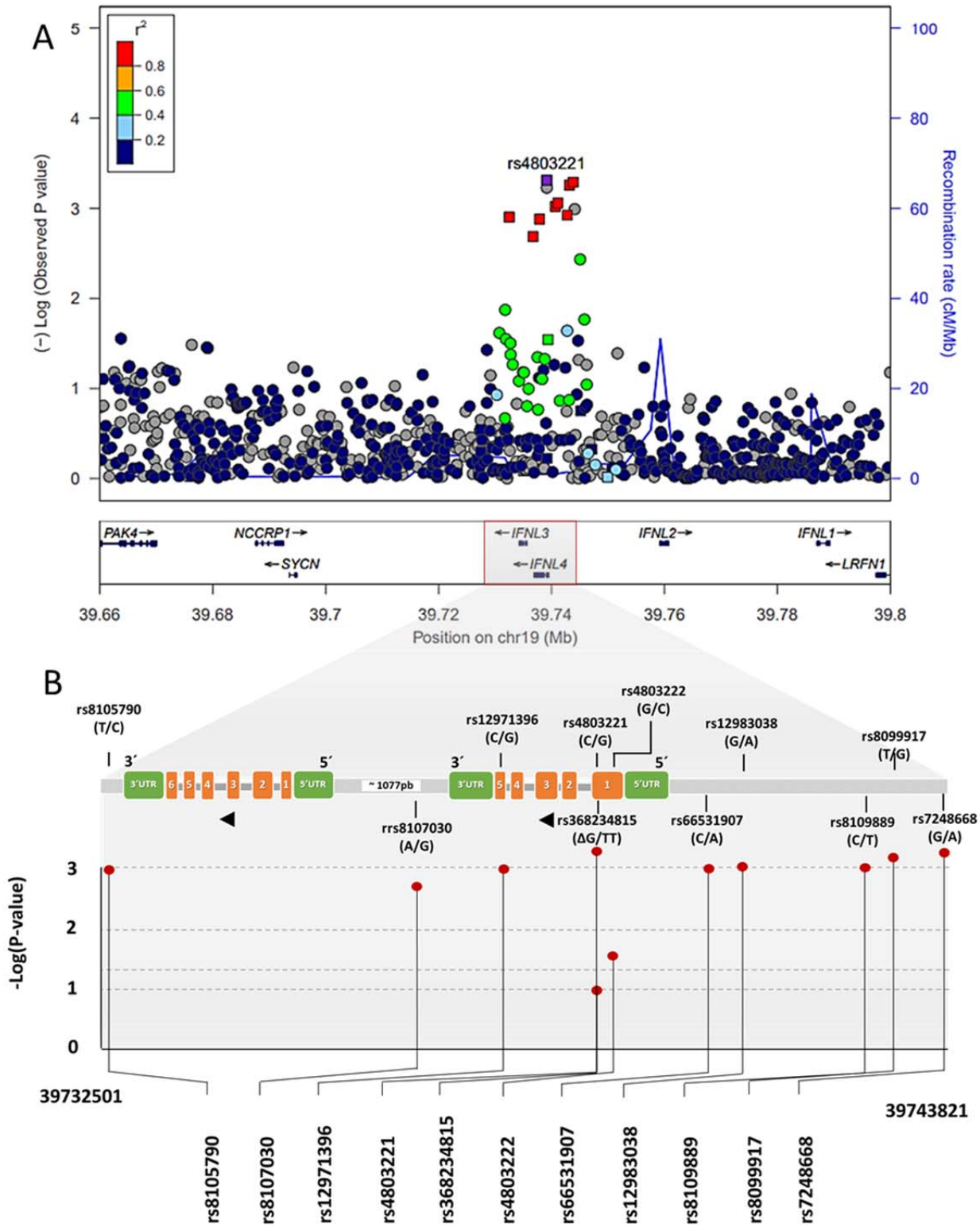
European Ancestry (Number of haplotypes= 1087)					
Haplotype rs4803221- rs368234815			**Frequency in Clearance (Number of haplotypes= 310)**	**Frequency in Persistence (Number of haplotypes= 777)**	**OR (95% CI, P value)**
G-ΔG Haplotype	G	ΔG	0.53	0.66	1
C-ΔG Haplotype	C	ΔG	0.47	0.34	1.7 (1.3-2.29,3.6x 10-5)
Haplotype rs368234815- rs8099917					
ΔG-G Haplotype	ΔG	G	0.51	0.63	1
ΔG-T Haplotype	ΔG	T	0.49	0.37	1.6 (1.3-2.16,1.76x 10-4)
African Ancestry (Number of haplotypes= 2333)					
Haplotype rs4803221- rs368234815			**Frequency in Clearance (Number of haplotypes= 319)**	**Frequency in Persistence (Number of haplotypes= 1994**	**OR (95% CI, P value)**
G-ΔG Haplotype	G	ΔG	0.29	0.30	1
C-ΔG Haplotype	C	ΔG	0.71	0.71	1.02 (0.78-1.32, 0.88)
Haplotype rs368234815- rs8099917					
ΔG-G Haplotype	ΔG	G	0.08	0.10	1
ΔG-T Haplotype	ΔG	T	0.92	0.90	1.25 (0.81-1.9, 0.29)


IFNL4 Gene
rs12979860
(C/T)
3´
5´
3'UTR
5
4
3
2
1
5'UTR
rs368234815
(TT/ΔG)
Exon 1
5'UTR
TT allele
TT/TT
Favorable
genotype
ΔG allele
ΔG/ΔG
Unfavorable
genotype
IFNλ4 protein
Effect associated on Hepatitis C virus infection
Hepatitis C virus infection
Clearance
Persistence

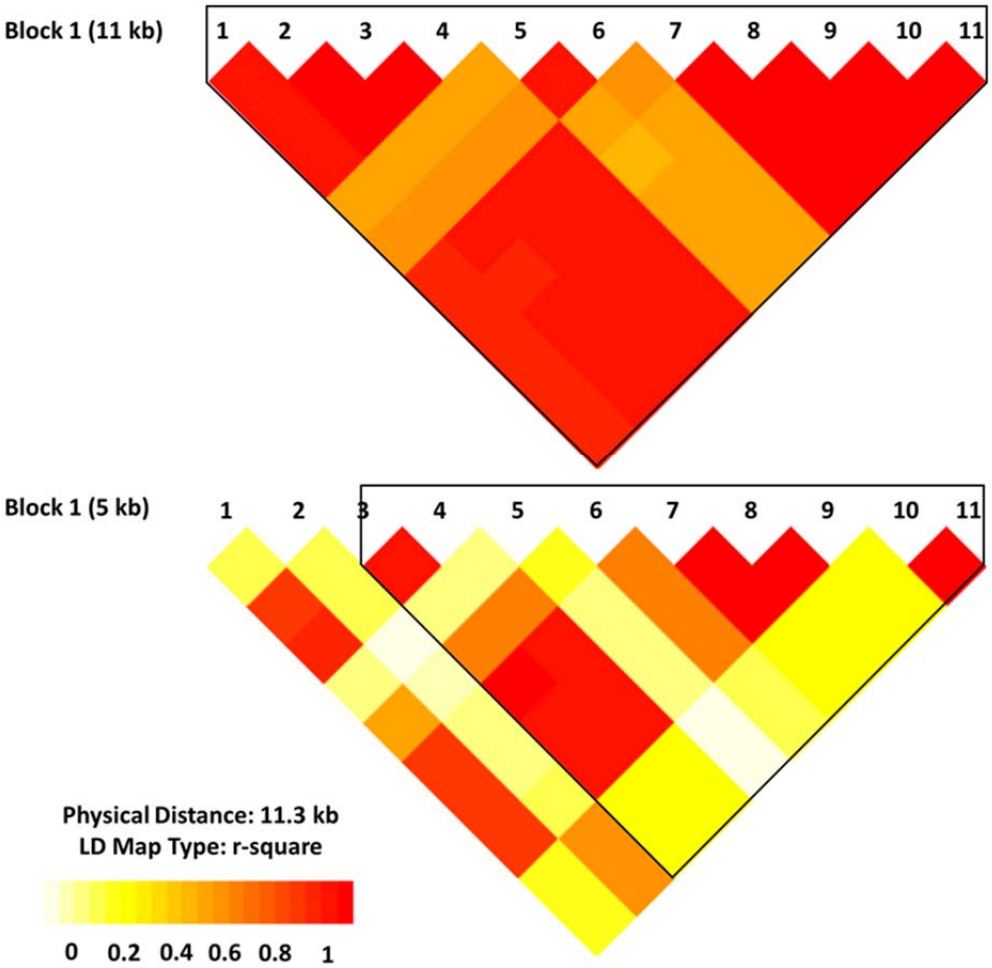

Panel of sequenced individuals
Panel
Genotype
Phenotype
rs12979860
(C/T)
3´
5´
rs368234815
(TT/ΔG)
Concordant
Discordant
Exon 1
5'UTR
TT/TT
(Favorable)
ΔG/ΔG
(Unfavorable)
Clearance
Persistance
N = 16
8 African ancestry
8 European ancestry
N = 15
13 African ancestry
2 European ancestry
N = 19
11 African ancestry
8 European ancestry
N = 14
13 African ancestry
1 European ancestry


Fine Mapping of Genome Wide Association Signal in IFNL locus
Analysis of sequenced panel of discordant and concordant individuals (n=19 European Ancestry, 45 African Ancestry)
Conditional analysis in independent imputed dataset (n=1717 European Ancestry, 1835 African Ancestry)
Identification of potential causal variants in all individuals (n=1739 European Ancestry, n= 1869 African Ancestry)
Variants with alternative allele count differences = > 5
Variants with significant effect after conditioning
Identification of credible set of potentially causal variants
Variants with evidence of independent effect in both approaches
Haplotype analysis. Identification of a core haplotype
99% Credible set of variants


chr19 (q13.2)
19p13.3
19p13.2
p13.11
19p12
19q12
13.11
13.12
19q13.2
13.32
q13.33
13.42
13.43
Scale
20 kb
hg19
chr19:
39,730,000
39,740,000
39,750,000
39,760,000
39,770,000
39,780,000
39,790,000
Amplified Segments
Fragment_A
Fragment_B
Fragment_C
Fragment_D
Fragment_E
Fragment_F
Fragment_G
Fragment_H
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
IFNL3
IFNL3
IFNL4
IFNL4
IFNL4
IFNL4
IFNL4
IFNL2
IFNL1
Main GWAS associated Variants In the IFNL region
rs12979860
rs368234815


A
r²
0.8
0.6
0.4
0.2
rs4803221
(−) Log (Observed P value)
Recombination rate (cM/Mb)
PAK4
NCCRP1
SYCN
IFNL3
IFNL4
IFNL2
IFNL1
LRFN1
39.66
39.68
39.7
39.72
39.74
39.76
39.78
39.8
Position on chr19 (Mb)
B
rs8105790 (T/C)
rs12971396 (C/G)
rs4803221 (C/G)
rs4803222 (G/C)
rs12983038 (G/A)
rs8099917 (T/G)
3′UTR
5′UTR
~ 1077pb
rrs8107030 (A/G)
rs368234815 (ΔG/TT)
rs66531907 (C/A)
rs8109889 (C/T)
rs7248668 (G/A)
-Log(P-value)
39732501
39743821
rs8105790
rs8107030
rs12971396
rs4803221
rs368234815
rs4803222
rs66531907
rs12983038
rs8109889
rs8099917
rs7248668

C


Block 1 (11 kb)
1 2 3 4 5 6 7 8 9 10 11
Block 1 (5 kb)
1 2 3 4 5 6 7 8 9 10 11
Physical Distance: 11.3 kb
LD Map Type: r-square
0 0.2 0.4 0.6 0.8 1