REVIEW

Pharmacogenomics of chronic hepatitis C therapy with genome-wide association studies

Chun-Hsiang Wang[1]
Yuchi Hwang[2]
Eugene Lin[2]

[1]Department of Hepatogastroenterology, Tainan Municipal Hospital, Tainan, Taiwan; [2]Vita Genomics, Inc., Taipei, Taiwan

Abstract: Chronic hepatitis C (CHC) is a liver disease characterized by infection with the hepatitis C virus (HCV) persisting for more than six months. Patients with CHC often stop pursuing the pegylated interferon (peg-IFN) and ribavirin (RBV) treatment because of the high cost and associated adverse effects. Therefore, it is highly desirable, both clinically and economically, to establish the determinants of response to distinguish responders from nonresponders, and to predict the possible outcomes of the peg-IFN and RBV treatments. The aim of this study was to review recent data on the pharmacogenomics of the drug efficacy of IFN in CHC patients. Single nucleotide polymorphisms (SNPs) can be used to understand the relationship between genetic inheritance and IFN therapeutic response. In the recent advent of scientific research, the genome-wide association study (GWAS), which is an alternative to the candidate-gene approach, is widely utilized to examine hundreds of thousands of SNPs by high-throughput genotyping technologies. In addition to the candidate-gene approach, the GWAS approach has recently been employed to study the determinants of HCV's response to therapy. Several recent findings have demonstrated that some SNPs in the interleukin 28B gene are closely associated with IFN responsiveness. These results promise to lead to mechanistic findings related to IFN responsiveness in this disease, and will probably have major contributions for individualized medicine and therapeutic decision making.

Keywords: chronic hepatitis C, genome-wide association study, interferon, pharmacogenomics, ribavirin, single nucleotide polymorphisms

Introduction

Chronic hepatitis C (CHC) affects more than 170 million individuals worldwide and is a chronic liver disease characterized by infection with the hepatitis C virus (HCV) persisting for more than six months.[1,2] Although combination therapy with pegylated interferon-alpha (peg-IFN-alpha) and ribavirin (RBV) has been the recommended treatment for CHC patients, many patients will not be cured by treatment.[1,2] In addition to limited efficacy, it is well known that some patients often stop completing therapy because of the high cost and significant adverse reactions.[1,2] Consequently, it would be highly desirable to identify the determinants of response to treatment that distinguish responders from nonresponders and predict the possible outcome of the peg-IFN-alpha and RBV treatment.[3,4]

The completion of the Human Genome Project has led to a new era of scientific research, including the revolutionary genome-wide association study (GWAS).[5] As an alternative to the candidate-gene approach, GWAS studies examine common genetic variations across the entire human genome in an attempt to identify genetic associations

Correspondence: Eugene Lin
Vita Genomics, Inc., 7 Fl., No. 6,
Sec. 1, Jung-Shing Road, Wugu Shiang,
Taipei, Taiwan
Tel +886 2 8976 9123 ext. 7751
Fax +886 2 8976 9523
Email eugene.lin@vitagenomics.com

73

with observable traits by utilizing high-throughput genotyping technologies to assay hundreds of thousands of single nucleotide polymorphisms (SNPs).[6,7] Unlike the candidate-gene approach, there is no a priori hypothesis about involved genes in the GWAS. A GWAS typically includes four parts: (1) selection of a large number of participants with the disease or trait of interest, and a suitable control group for comparison; (2) DNA isolation, array platforms, and data review to ensure high genotyping quality; (3) statistical tests for associations between SNPs and the disease and/ or trait; and (4) replication of identified associations in an independent population or experiment of functional studies.[8] In a typical GWAS, frequencies of 500,000 to 2 million SNPs are compared in cases and controls. This GWAS approach presents several challenges, such as the issue of multiple comparisons and how to account for ethnic and geographic differences in the prevalence of SNPs.[9,10] The situation of multiple comparisons refers to the probability that SNP frequencies in cases and controls would differ simply because chance naturally increases when there are hundreds of thousands of comparisons. The issue is usually resolved by requiring a very large sample size or by detecting a SNP with a very strong association.[11,12]

In this paper, we review the pharmacogenomics of the drug efficacy of interferon (IFN) in CHC patients. In addition to the candidate-gene approach, the GWAS approach has recently been employed to study the determinants of HCV's response to therapy. Several recent studies have demonstrated that some SNPs in the interleukin 28B (IL28B) gene are closely associated with IFN-responsiveness.[13–18] The IL28B gene codes for IFN-lambda-3, which is one of the recently discovered type III interferons, and is ubiquitous in vertebrates.[19] Unlike the single-exon type I interferons, IFN-lambda-3 is typically encoded by five or six exons.[19]

Genome-wide association studies
GWAS by Ge and colleagues

In a GWAS, Ge and colleagues studied genetic variation as a predictor of hepatitis C treatment-induced viral clearance.[13] They genotyped 1,671 patients from the Individualized Dosing Efficacy versus Flat Dosing to Assess Optimal Pegylated Interferon Therapy (IDEAL) trial using the Human 610-quad BeadChip® (Illumina Inc., San Diego, CA, USA).[13] Treatment-naive US patients with HCV genotype 1 underwent 48 weeks of therapy with peg-IFN-alpha-2b or peg-IFN-alpha-2a combined with RBV, and were assessed for a sustained virologic response (SVR), defined as undetectable levels of HCV RNA at 24 weeks after treatment ended.[13]

In this GWAS, Ge and colleagues reported an SNP, rs12979860, in the IL28B gene, that was strongly associated with SVR in all patient groups within the study populations, including European Americans ($P = 1.06 \times 10^{-25}$), African Americans ($P = 2.06 \times 10^{-3}$), and Hispanics ($P = 4.39 \times 10^{-3}$).[13] The SNP rs12979860 resides 3 kb upstream of the IL28B gene on chromosome 19. The CC genotype of the IL28B rs12979860 SNP was found to be associated with higher SVR rates than the TT or TC genotypes across all population groups.[13] The magnitude of this association between the CC genotype and SVR was greater than that of other clinical factors currently utilized to predict treatment response in patients infected with genotype 1 HCV, including baseline viral load, baseline fibrosis, and ethnicity (odds ratio [OR] range: 5.6–7.3 for CC genotype, 2.4–5.1 for viral load, 1.1–4.1 for fibrosis, and 3.1 for ethnicity, respectively).[13] The SVR rates across different population groups were found to have a substantial concordance with C allele frequency at IL28B rs12979860. Thus, the difference in frequency of the C allele can be estimated to explain approximately half of the difference in treatment response rates between African American and European American patients.[13]

Furthermore, Ge and colleagues suggested that the C allele frequency of the IL28B rs12979860 variant was significantly decreased in chronically infected patients compared with ethnically matched controls (0.63 versus 0.73; $P < 2.5 \times 10^{-6}$), indicating an association between the C allele and a higher rate of natural clearance of HCV.[13] However, they did not perform comparison between matched patients known to have, or not to have, naturally cleared the HCV infection.

Finally, Ge and colleagues sequenced the IL28B gene in 96 of the individuals they genotyped, and found two variants in linkage disequilibrium with rs12979860, including a G > C transition 27 bp upstream of the translation initiation codon (rs28416813), and a nonsynonymous coding SNP (213A > G, K70R, rs8103142; $r^2 > 0.85$).[13] However, the investigators did not pursue functional analysis in their study. Owing to the high degree of correlation among the three SNPs, tests for independence were not possible.[13] Thus, it could not be determined which, if any, of the SNPs is uniquely responsible for the association with SVR.

GWAS by Suppiah and colleagues

In a similar GWAS for HCV's response to therapy, Suppiah and colleagues studied an Australian population of northern European ancestry infected with HCV genotype 1, including 162 nonresponders to peg-IFN-alpha and RBV,

and 131 responders.[14] In the first stage, they utilized the Infinium™ HumanHap300 (Illumina Inc.) or the CNV370-Quad genotyping BeadChip (Illumina Inc.) with data for 311,159 SNPs. The Cochrane–Armitage trend test was employed to test the association between genotypes and response status. The threshold for genome-wide significant association was set at $P < 1.6 \times 10^{-7}$ (0.05/311,159).[14] One SNP, rs8099917, in the intergenic region between IL28A and IL28B, was shown to be associated with response status at genome-wide significance ($P = 7.06 \times 10^{-8}$; OR = 3.36; 95% CI: 2.15–5.35).[14] They then replicated their results in a larger independent cohort of Europeans from the UK, Germany, Italy, and Australia. The association between SVR and the G allele of the SNP rs8099917 was found to reach genome-wide significance in the overall analysis of the first-stage and replication cohorts ($P = 9.25 \times 10^{-9}$; OR = 1.98; 95% CI: 1.57–2.52).[14] Compared with noncarriers, heterozygous and homozygous carriers of the G allele of rs8099917 had ORs of 1.64 and 2.39 for SVR, respectively.[14] Several other SNPs were also reported to have suggestive associations with a P value of less than 10^{-4}.

Moreover, Suppiah and colleagues genotyped a total of 20 tag SNPs in the IL28B, IL28A, and IL29 gene cluster (on chromosome 19: 44389814–44480955), and identified a six-allele haplotype (GCCTAG) comprised of rs12980275, rs8105790, rs8103142, rs1085327, rs8109886, and rs8099917.[14] The rs8099917 SNP tags this haplotype, which splits responders and nonresponders ($P = 3.03 \times 10^{-9}$; OR = 2.0; 95% CI: 1.58–2.50).[14] This haplotype was shown to be present in 31.5% of nonresponders compared with 18.8% of responders.[14] To verify whether this haplotype might affect the expression of both IL28A and IL28B genes, they tested the constitutive expression of these genes in whole blood in 49 healthy participants, and observed that the non-responder haplotype had lower expression of IL28A and IL28B ($P = 0.044$).[14] In addition, Suppiah and colleagues found that the G allele of rs8099917 was able to predict non-response with 57% sensitivity and 63% specificity, based on a model of dominant inheritance.[14] Therefore, this SNP, in combination with other parameters and genetic variants, should be useful for clinical management.

GWAS by Tanaka and colleagues

Similarly, in the third GWAS for the drug efficacy of peg-IFN-alpha and RBV, Tanaka and colleagues conducted a multi-center study in Japanese patients with HCV genotype 1.[15] They utilized the Affymetrix® (Affymetrix, Santa Clara, CA, USA) SNP 6.0 genome-wide SNP typing array for 900,000 SNPs.

In the first phase of the study, there were a total of 78 patients with null virologic response (NVR) (less than a 2 log unit decline in serum HCV RNA at 12 weeks, and detectable viremia at 24 weeks of treatment), and 64 patients with a virologic response (VR), achievement of SVR, or transient VR.[15] The patient cohort was treated with either standard-dose peg-IFN-alpha-2b or peg-IFN-alpha-2a, in combination with ribavirin, and maintained adherence of more than 80% dose for both drugs during the first 12 weeks of treatment.[15]

Using a minor allele dominant model, Tanaka and colleagues revealed that two SNPs, including rs8099917 ($P = 3.11 \times 10^{-15}$) and rs12980275 ($P = 1.93 \times 10^{-13}$), which locate close to IL28B on chromosome 19, were in strong association with NVR to peg-IFN-alpha with RBV treatment.[15] They also observed that the frequencies of minor allele-positive patients were much higher in the NVR group than in the VR group for both SNPs (74.3% in NVR and 12.5% in VR for rs12980275; 75.6% in NVR and 9.4% in VR for rs8099917), and individuals homozygous for the minor allele were in only the NVR group.[15] It is interesting to point out that Tanaka and colleagues identified the same SNP, rs8099917, as in the study by Suppiah and colleagues,[14] to be in association with treatment response. Another result similar to the study of Suppiah and colleagues[14] was that they demonstrated lower IL28B mRNA expression in peripheral blood mononuclear cells in individuals carrying the minor alleles in their study.[15]

In the replication phase of the study, Tanaka and colleagues further identified 16 SNPs by analyzing the 40 kb region that contained the rs8099917 and rs12980275 SNPs.[15] They genotyped an independent set of 172 Japanese patients (50 NVR and 122 VR), together with the patient samples of the first phase, using the DigiTag2 assay.[15] In the replication analysis, they confirmed the initial association with combined p values and ORs for NVR versus transient VR (rs8099917: $P = 1.11 \times 10^{-27}$; OR = 27.2; and rs12980275: $P = 3.99 \times 10^{-24}$; OR = 18.5).[15] Furthermore, they identified six SNPs with significant association within a 15.7 kb region that included IL28B, and then further sequenced the region to identify three more. In all, they observed that the strongest associations with NVR were in seven SNPs (including rs8105790, rs11881222, rs8103142, rs28416813, rs4803219, rs8099917, and rs7248668) that are located in the downstream flanking region, the third intron, the third exon, the first intron, and the upstream flanking region of the IL28B gene.[15] Using logistic regression analysis, they identified the SNP rs8099917 as the most significant factor for predicting NVR (OR = 37.68).[15]

GWAS by Rauch and colleagues

Very recently, Rauch and colleagues performed the fourth GWAS of a relatively genetically homogenous population to screen for host genetic determinants of HCV persistence and response to therapy.[16] In their study, there were 1,362 Caucasian Swiss patients, including 1,015 subjects with persistent HCV infection, and 347 subjects with spontaneous HCV resolution (where 448 subjects were co-infected with human immunodeficiency virus).[16] With multivariate logistic regression, they assessed responses to peg-IFN-alpha and RBV in 465 individuals, using more than 500,000 SNPs.[16]

Rauch and colleagues found that CHC was associated with 7 SNPs in the IL28B locus.[16] The strongest association with spontaneous recovery was detected for rs8099917, a T/G SNP whose distribution was exceedingly unlikely to have occurred by chance (OR = 2.31; 95% CI: 1.74–3.06; $P = 6.07 \times 10^{-9}$).[16] Furthermore, the frequency of the minor G allele at rs8099917 was over-represented in the subset of those with chronic infection who failed to respond to peg-IFN-alpha and RBV, compared with those who achieved SVR (OR = 5.19; 95% CI: 2.90–9.30; $P = 3.11 \times 10^{-8}$), with the strongest effects in patients with HCV genotype 1 or 4.[16]

Racial differences

As mentioned previously, four groups have found pieces of the HCV outcome puzzle near the IL28B gene.[13–16] Three of them found that the T allele in rs8099917 was associated with SVR to treatment.[14–16] On the other hand, one group observed that persons who were homozygous for the major C (versus T) allele at the SNP rs12979860 near the IL28B gene were 2-fold more likely to respond to treatment than those homozygous for the alternative nucleoside, T.[13] It is interesting to mention that the SNPs rs8099917 and rs12979860 are just 4,378 bases apart.[16] In addition, their linkage was high (D' = 0.98; $r^2 = 0.5$) among the Swiss population in the GWAS by Rauch and colleagues.[16] Similarly, Ge and colleagues reported high linkage between these two SNPs in Caucasians ($r^2 = 0.52$) but very low linkage in African Americans ($r^2 = 0.07$).[13] We can utilize the HapMap database[20] to provide a comparison among the five populations (African ancestry in Southwest USA [ASW], Utah residents with Northern and Western European ancestry from the CEPH collection [CEU]; Han Chinese in Beijing, China [CHB]; Japanese in Tokyo, Japan [JPT]; Yoruba in Ibadan, Nigeria [YRI]) in terms of allele frequencies. Interestingly, data from the HapMap database show that the risk allele (G) in rs8099917 is infrequent in the CHB and African populations (3% in CHB and 2% in YRI, respectively), which is

lower than the 18% in the CEU.[20] Because patients with African ancestry are more likely to have viral persistence and less likely to respond to treatment, this allele would not explain the racial differences in HCV outcomes.[21] In contrast, the risk allele (T) in rs12979860 is more common in blacks than in Caucasians (6.6% in CHB and 69.2% in YRI in HapMap, respectively).[20] Thus, further research is needed to determine whether this IL28B gene explains racial differences in the outcome of this global infectious disease.[21]

Type III interferons

As presented earlier, at least four independent GWAS studies provide significant evidence for the role of the IL28B gene in the pathogenesis of HCV infection.[13–16] However, still more pieces of data are needed to complete the mechanistic picture.[21] The IL28B gene encodes IFN-lambda-3, which is one of the recently discovered type III interferons (or lambda interferons), and belongs to the IL10 superfamily.[22] Other type III interferons are the cytokine IL28A and IL29. Type III interferons are induced through mechanisms similar to, and probably synergistic with, type I interferons (such as IFN-alpha), although they utilize distinct receptor complexes to ultimately trigger similar intracellular signaling programs (including phosphorylation of the kinases JAK1 and TYK2, activation of the transcription factor complex containing STAT1, STAT2, and interferon regulatory factor 9, and upregulation of a similar set of genes that share the IFN-stimulated response element promoter sequence).[23,24]

Some murine studies have revealed the preferential expression of IFN-lambda receptors on epithelial surfaces, and have suggested that type III interferons may allow the host to rapidly eliminate viruses at the major portals of entry into the body before infection is established, without activating other arms of the immune system.[25–27] Nonetheless, one important difference established between the murine and human systems is that the interleukin 28 receptor alpha chain (receptor complex induced by IFN-lambda) is expressed in human hepatocytes, whereas the murine liver seems unlikely to respond to IFN-lambda.[25–27] This finding suggests that IFN-lambda contributes to host defense against hepatotrophic viruses, such as HCV, in humans.[28] Like type I interferons, lambda interferons have activity against HCV and other viral infections *in vitro* and *in vivo*.[29] However, *in vitro* exogenous IFN-lambda induces a slower, more sustained abundance of IFN-stimulated genes than IFN-alpha.[29] Although these findings answer why IFN-lambda might play a key role in HCV recovery, they do not explain why treatment-associated resolution of HCV infection is associated with certain base

sequences located upstream of the start codon for the cytokine IL28B.[21]

IL28B expression

The recent GWAS articles linking the IL28B genotype to IFN-alpha therapeutic response have triggered intensive research to establish underlying mechanisms for the association.[28] At this point, there are very few possible explanations and, few strong, mechanistic clues.[21]

It is speculated that these variants correlate with the regulation of the cytokine IL28B transcription, because these SNPs are located upstream of the IL28B gene.[21] The observations cited earlier by Suppiah and colleagues, and Tanaka and colleagues strongly suggest that the identified SNPs do indeed alter the expression of the IL28B and, perhaps, IL28A genes.[14,15] These two studies found that those who carried the G risk allele at the SNP rs8099917 had lower mRNA expression of the cytokine IL28B in peripheral blood mononuclear cells.[14,15] On the other hand, Ge and colleagues reported no difference in cytokine IL28B expression in peripheral blood mononuclear cells from 80 HCV-uninfected persons homozygous for a proxy allele for the SNP rs12979860, using the SNPExpress database.[13]

However, it remains unknown how such altered expression of IFN-lambda would lead to reduced responsiveness to exogenously administered IFN-alpha.[28] It is likely that IFN-lambda may play a role in enhancing the effect of IFN-alpha signaling pathways.[14,15] Moreover, it is plausible to explain the association between IL28B and response by the following consistent observation made by several investigators: the increased expression of IFN-stimulated genes at baseline in patients who were destined to fail to respond to IFN-alpha.[30–32] It has been shown that upregulation in a specific set of eight IFN-stimulated genes predicts nonresponse to exogenous peg-IFN therapy.[30] Furthermore, it has been reported recently that regulation differs in the infected tissue, and even between cell types within the liver, for some IFN-stimulated genes.[31] It has also been demonstrated that a strong upregulation of IFN-stimulated genes was induced in the livers of patients without rapid VR to peg-IFN, compared with low baseline expression and high post-treatment expression in those with rapid VR.[32]

Similar to the protective effect of the least inhibitory interactions between killer cell immunoglobulin-like receptor (KIR) and HLA-C, it is likely that the cytokine IL28B produced in persons with the protective allele diminishes cytokine IL13 to a lesser extent than the molecule with the risk allele.[33,34] It is also possible that these SNPs are markers

of another DNA sequence that modifies the cytokine IL28B.[21] In addition, both rs12979860 and rs8099917 variants are strongly associated with a nonsynonymous IL28B mutation, due to the fact that the favorable SNP haplotype usually occurs in association with an arginine instead of a lysine at position 70 (rs8103142).[13,17,21] This change is distant from the putative receptor binding location, and its mechanistic significance remains unclear.[35]

Practical implications

In addition to the mechanistic questions of profound importance, the recent GWAS findings on genetic variation as a predictor of outcome in patients treated for HCV infection will have major practical implications.[13–16] Several recent GWAS studies have demonstrated remarkable associations between SNPs near or within the region of the IL28B gene, which codes for IFN-lambda-3.[13–16] These results promise to lead to important mechanistic findings related to IFN responsiveness in this disease, and will probably have major contributions for individualized medicine and therapeutic decision making.[13–16]

Ge and colleagues suggest that, in the near future, advance knowledge of the genotype of patients infected with HCV could become an important component of the clinical decision to initiate treatment.[13] These findings may also lead to more individualized treatment regimens with regard to both the chosen medicines and the duration of therapy.[13–16] Moreover, the critical importance of the IL28B gene region in mediating response has emphasized interest in the development of IFN-lambda as a therapeutic agent for patients with CHC.[36–38] For instance, the early studies of a pegylated form of IL29 (IFN-lambda-1) have demonstrated promising antiviral activity with the potential benefit of reduced hematologic toxicity, owing to the hepatocyte-specific receptor profile for IFN-lambda (versus IFN-alpha).[36–38]

Relevant candidate-gene studies
HCV natural clearance

As mentioned previously, Ge and colleagues also demonstrated that the C allele frequency of the IL28B rs12979860 SNP was significantly reduced in the chronically infected cohort, compared with ethnically matched controls (0.63 versus 0.73; $P < 2.5 \times 10^{-6}$), suggesting an association between the C allele and a higher rate of natural clearance of HCV.[13] In a subsequent candidate-gene study, Thomas and colleagues confirmed this association by genotyping this IL28B rs12979860 variant in HCV cohorts comprised of individuals who spontaneously cleared the virus (n = 388) or had

persistent infection (n = 620).[17] They showed that the C allele occurred more frequently among patients who spontaneously cleared HCV than those who did not, among both European and African ancestry (80.3% versus 66.7%; $P = 7 \times 10^{-8}$, and 56.2% versus 37%; $P = 1 \times 10^{-5}$, respectively).[17] Furthermore, they observed that patients with the CC genotype at IL28B rs12979860 were three times more likely to have natural clearance of HCV compared with patients with the combined CT and TT genotypes (OR = 0.33; $P < 1 \times 10^{-12}$).[17] Finally, it is likely that the 213A > G change (in the SNP rs8103142) alters the function of the IFN-lambda-3, explains the difference in spontaneous HCV clearance after acute infection, and mediates treatment response to IFN.[17]

Replicated association

As discussed previously, a GWAS study by Ge and colleagues has linked response to peg-IFN and RBV with a common rs12979860 SNP in the IL28B gene on chromosome 19.[13] McCarthy and colleagues conducted a replication candidate-gene study to investigate the association between the SNP rs12979860 and treatment response in a diverse cohort of chronic HCV patients.[18] A cross-sectional study was performed using data from 1,021 consecutive patients enrolled in the Duke Hepatology Clinic Research Database and Biorepository.[18] They analyzed DNA, clinical, and demographic data, along with validated data of the response of 231 subjects to peg-IFN and RBV. The study included Caucasians (n = 178), African Americans (n = 53), and HCV genotypes 1 (n = 186), and 2/3 (n = 45). The IL28B rs12979860 genotype was tested for an association with SVR.

McCarthy and colleagues reported that the rs12979860 CC genotype (found in approximately 40% of Caucasians) predicted a SVR to therapy among Caucasians (OR = 5.79; 95% CI: 2.67–12.57; $P = 9.0 \times 10^{-6}$) independent of HCV genotype and other covariates.[18] In addition, the rs12979860 CC genotype predicted a SVR with 78% specificity and 65% sensitivity in patients infected with HCV genotype 1, which was better than the HCV genotype currently used to predict treatment response.[18] Thus, McCarthy and colleagues concluded that the IL28B rs12979860 genotype is a significant predictor that might be used to determine the best course of peg-IFN and RBV treatment for patients with chronic HCV infection.[18]

IFN alpha pathway

In a recent candidate-gene study, Welzel and colleagues investigated the association between genetics and HCV treatment response with a focus on the IFN-alpha pathway.[39]

They genotyped 56 SNPs along the IFN-alpha pathway in 1,051 patients in the Hepatitis C Long-term Treatment Against Cirrhosis (HALT-C) trial, utilized TaqMan® assays (Applied Biosystems, Carlsbad, CA, USA) with analysis on the ABI 7900HT platform (Applied Biosystems), and focused on European Americans (n = 581) for purposes of statistical power.[39] Participants had advanced chronic hepatitis C (Ishak fibrosis score ≥3), had not previously responded to IFN treatment with or without RBV, had a Child–Turcotte–Pugh score of less than seven, and were treated with peg-IFN and RBV for 24 weeks, and then up to 48 weeks if undetectable at week 20.[39]

In the study by Welzel and colleagues, they examined the SNPs in genes encoding IFN-alpha, the IFN-alpha receptor, JAK/tyrosine kinase/STAT 1 and 2, all part of the signal transduction via the JAK-STAT pathway, IFN-alpha induced genes with antiviral properties, including adenosine deaminase/eukaryotic translation initiation factor 2A-alpha kinase 2/NFKB1/myxovirus resistance 1/2´5´-oligoadenylate synthetase 1, as well as interferon regulatory factor.[39]

After using logistic regression to control for the effect of other variables associated with treatment response, associations were reported among 581 European American patients as follows: the minor variant for the intronic IFN-alpha receptor 1 (IFNAR1) IVS1–22T > G (rs2243592) SNP occurred at lower frequency in patients with SVR (adjusted OR = 0.57; $P = 0.02$); the C variant for the IFN-alpha receptor 2 (IFNAR2) EX2–33T > C (rs4986956) SNP occurred at higher frequency in patients with SVR (adjusted OR = 2.09; $P = 0.02$); Janus kinase 1 (JAK1) IVS22+112G > T (rs2254002) was associated with SVR (adjusted OR = 1.66; $P = 0.04$); and RNA-specific adenosine deaminase (ADAR) EX9+14G > A (rs1127309) occurred at higher frequency in patients with SVR (adjusted OR = 1.67; $P = 0.03$).[39] In addition, the A allele of promoter region 2256 (rs12720217) in the tyrosine kinase 2 (TYK2) gene was found at higher frequency in European Americans with SVR (adjusted OR = 1.51; $P = 0.05$).[39] Moreover, a strong relationship was supported in the African American patients, such that all 10 patients with SVR who carried the TYK2–2256A variant, compared with 57% of nonresponders ($P = 0.006$).[39]

Therefore, Welzel and colleagues established the role of the gene products along the IFN-alpha pathway, and concluded that their most promising result was this variant, TYK2–2256A, associated with treatment response.[39] One limitation of the study is that functional analysis was beyond the scope of their study. Second, these findings may not be suitably generalized to other populations. Although the type III

interferons are thought to signal through similar mechanisms to the type I interferons/IFN-alpha, the associations identified by Welzel and colleagues in JAK1 (rs2254002), TYK2 (rs12720217), and ADAR (rs1127309) were not found in the aforementioned GWAS studies.[28] This may be owing to the fact that the population in the HALT-C trial represents a unique population with advanced disease and previous failure to respond to therapy.[39]

Modeling HCV treatment response

The efficacy of IFN-alpha and RBV is likely influenced by the combined effects of a number of genetic variants.[3,4,40] Accumulating evidence reveals that SNPs could be used as genetic markers to predict IFN-alpha and RBV treatment outcome in CHC.[3,4,40] Several studies[41–43] in different populations support the implication that the effects of IFN-alpha and RBV are associated with genetic variants. In addition, the genetic differences have been analyzed and found to be associated with IFN-alpha and RBV responses, using a multiple logistic regression method.[3]

Artificial neural network (ANN) algorithms are generally adopted for complex classification applications because of the advantages of ANN algorithms, such as nonlinearity, fault tolerance, universality, and real-time operation.[44,45] ANN algorithms have been employed to build a prediction model for the drug efficacy of IFN-alpha and RBV in CHC patients based on SNPs and other clinical factors.[4,40] Moreover, the possible nonlinear relationships between genetic variants and antidepressant response have been explored using ANN algorithms in pharmacogenomic studies.[46,47]

Several previous findings[3,4,40] reported modeling IFN-alpha and RBV treatment response by using logistic regression or ANN methods without feature selection. In a recent study, Ke and colleagues extended the previous research and applied both ANN algorithms and logistic regression with feature selection to predict IFN-alpha and RBV treatment outcomes using genetic factors.[48] The cohort of 523 CHC patients was original to the previous study by Lin and colleagues.[4] Briefly, blood samples were collected from 523 CHC patients at National Taiwan University Hospital, Kaohsiung Medical University Hospital, Kaohsiung Chang-Gung Memorial Hospital, and Tri-Service General Hospital in Taiwan from 2002 to 2004. Patients whose serum HCV RNA became negative and lasted for more than 6 months after the end of treatment were defined as SVRs of the treatment. Those who still remained viremic were defined as non-responders. There were 523 patients, including 350 SVRs and 173 nonresponders.[4] They further converted the clinical diagnostic data into numerical forms (ie, 1 for "SVR" and 0 for "nonresponder", respectively). Furthermore, genomic DNA was amplified using a commercially available SNP detection kit (Vita Genomics, Inc., Taiwan) according to the manufacturer's instructions.[4,40] The SNP genetic markers of the subjects were generated at the high-throughput genomics lab of Vita Genomics, Inc. In that study, they only focused on 24 SNPs.[49] Furthermore, they used two families of classification algorithms, including multilayer feedforward neural network (MFNN) and logistic regression, as a basis for comparisons. An MFNN is one type of ANN model where connections between the units do not form a directed cycle.[50] These classifiers were performed using the Waikato Environment for Knowledge Analysis (WEKA) software (University of Waikato, Hamilton, New Zealand).[51] In the learning phase of this MFNN, the back-propagation algorithm[52] is employed for the learning scheme.

To identify a subset of clinical factors that maximizes the performance of the prediction model, Ke and colleagues employed the wrapper-based feature selection approach, where the feature selection algorithm acts as a wrapper around the classification algorithm.[53] The wrapper-based approach conducts best-first search for a good subset and uses the classification algorithm itself as part of the function for evaluating feature subsets.[54] They adopted the area under the receiver operating characteristic curve (AUC) for evaluating predictive ability of classifiers owing to the fact that AUC is a better performance metric for accuracy.[55] The AUC of a classifier can be interpreted as the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative one.[55] The higher the AUC, the better the learner.[56]

This work by Ke and colleagues is the first study that proposes to use MFNN and logistic regression with the wrapper-based feature selection method to model the drug responding status in CHC patients using genetic factors.[48] They developed a pharmacogenomic methodology to predict the drug efficacy of IFN-alpha and RBV in CHC patients based on genetic factors such as SNPs. Their results demonstrate that a trained MFNN model is a promising method for providing the inference from genetic factors, such as SNPs, to the responsiveness of IFN-alpha and RBV. Their findings also suggest that their tool may provide a medical reference prior to treatment, based on the information of genetic factors such as SNP genotypes. In addition, the wrapper-based method has the advantage that it includes the interaction between the feature subset search and the classification model.[53] However, the wrapper-based method may have a risk of over-fitting.[53,57]

A similar study by Lin and colleagues[4] has reported utilizing MFNN algorithms to evaluate the possible nonlinear interactions between IFN-alpha and RBV response, and factors such as seven SNPs, viral genotype, viral load, age, and gender. The same cohort of 523 patients with CHC was used in both studies. Ke and colleagues demonstrated that an MFNN network with one hidden layer had an accuracy of 80.4%.[48] On the other hand, Lin and colleagues reported that an MFNN network with one hidden layer had an accuracy of 77.4%.[4] The difference between these two studies was that the study by Ke and colleagues used 24 SNPs instead of only seven polymorphisms.[4,48] Moreover, Lin and colleagues did not employ the wrapper-based feature selection method. As shown in the simulation results of Ke and colleagues, their MFNN prediction model performed better than the one by Lin and colleagues in terms of accuracy.[4,48] These preliminary results suggested that an MFNN model may be considered as a good method to deal with the complex nonlinear relationship between clinical factors and the responsiveness of IFN-alpha and RBV. Unlike logistic regression, MFNN has the ability to model the multidimensional and nonlinear relationships between variables, as found in complex medical applications.[58–60] Their simulation results have validated the implication that MFNNs, with only one hidden layer, should be adequate as universal approximators of any nonlinear functions.[61–63]

Conclusions

In this study, we reviewed several recent GWAS findings and relevant studies for the drug efficacy of IFN in CHC patients. Although further research is necessary to study the functional mechanism underlying the association between the IL28B gene and response, these related studies provide incontrovertible genetic evidence for a role of the IL28B gene in spontaneous and treatment-related recovery from HCV infection.[13–18] In the meantime, it is clear that the genotype CC at the SNP rs12979860, 3 kb upstream of the IL28B gene, is significantly associated with response to peg-IFN and RBV for patients with chronic genotype 1 infection, as well as natural clearance, and that the presence of the minor allele G at rs8099917 is significantly associated with virologic nonresponse.[13–18] This work also underscores the power of large-scale genetic studies, such as GWAS, to investigate a greater diversity of populations in the clinical expression of global infectious diseases and their treatments.[21] In addition, although future research is needed, there is a definite suggestion of association between the A allele of the SNP rs12720217 in the TYK2 gene and SVR in African American patients.[39]

Now we have a major new piece in the puzzle, after some pieces fitting the puzzle have been examined.[21] To improve and personalize HCV treatment and prevention worldwide, future effort will have to link these findings to other pieces, until the picture of HCV resolution is sufficiently clear.[21] Furthermore, these findings suggest that modeling tools may allow patients and doctors to make more informed decisions based on clinical factors such as SNP genotyping data.[4,47–49] Over the next few years, genetic tests for pretreatment prediction may become a reality in patient care, after prospective large clinical trials to validate clinical factors and genetic markers.[4,40] This may also provide potential drug targets for the development of alternative therapeutic agents to treat CHC patients, especially for those nonresponders.[4,40]

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Lo Re V 3rd, Kostman JR. Management of chronic hepatitis C. *Postgrad Med J*. 2005;81(956):376–382.
2. Modi AA, Liang TJ. Hepatitis C: a clinical review. *Oral Dis*. 2008;14(1): 10–14.
3. Hwang Y, Chen EY, Gu ZJ, et al. Genetic predisposition of responsiveness to therapy for chronic hepatitis C. *Pharmacogenomics*. 2006;7(5): 697–709.
4. Lin E, Hwang Y, Wang SC, Gu ZJ, Chen EY. An artificial neural network approach to the drug efficacy of interferon treatments. *Pharmacogenomics*. 2006;7(7):1017–1024.
5. Johnson AD, O'Donnell CJ. An open access database of genome-wide association results. *BMC Med Genet*. 2009;10:6.
6. Christensen K, Murray JC. What genome-wide association studies can do for medicine. *N Engl J Med*. 2007;356(11):1094–1097.
7. Need AC, Goldstein DB. Whole genome association studies in complex diseases: where do we stand? *Dialogues Clin Neurosci*. 2010;12(1): 37–46.
8. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008;299(11):1335–1344.
9. Goldstein DB. Common genetic variation and human traits. *N Engl J Med*. 2009;360:1696–1698.
10. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. *Nat Rev Genet*. 2010;11(5):356–366.
11. Elbers CC, van Eijk KR, Franke L, et al. Using genome-wide pathway analysis to unravel the etiology of complex diseases. *Genet Epidemiol*. 2009;33(5):419–431.

12. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet*. 2010;86:6–22.

13. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature*. 2009;461: 399–401.

14. Suppiah V, Moldovan M, Ahlenstiel G, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet*. 2009;41(10):1100–1104.

15. Tanaka Y, Nishida N, Sugiyama M, et al. Genome-wide association of IL28B with response to pegylated interferon-a and ribavirin therapy for chronic hepatitis C. *Nat Genet*. 2009;10:1105–1111.

16. Rauch A, Kutalik Z, Descombes P, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology*. 2010;138:1338–1345.

17. Thomas DL, Thio CL, Martin MP, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature*. 2009;461: 798–802.

18. McCarthy JJ, Li JH, Thompson A, et al. Replicated association between an IL28B gene variant and a sustained response to pegylated interferon and ribavirin. *Gastroenterology*. In press 2010.

19. Fox BA, Sheppard PO, O'Hara PJ. The role of genomic data in the discovery, annotation and evolutionary interpretation of the interferon-lambda family. *PLoS One*. 2009;4:e4933.

20. The International HapMap Consortium. The International HapMap Project. *Nature*. 2003;426:789–794.

21. Thio CL, Thomas DL. Interleukin-28b: a key piece of the hepatitis C virus recovery puzzle. *Gastroenterology*. 2010;138(4):1240–1243.

22. Dellgren C, Gad HH, Hamming OJ, Melchjorsen J, Hartmann R. Human interferon-lambda3 is a potent member of the type III interferon family. *Genes Immun*. 2009;10(2):125–131.

23. Larkin J, Jin L, Farmen M, et al. Synergistic antiviral activity of human interferon combinations in the hepatitis C virus replicon system. *J Interferon Cytokine Res*. 2003;23:247–257.

24. Szabo G, Dolganiuc A. Hepatitis C and innate immunity: recent advances. *Clin Liver Dis*. 2008;12(3):675–692.

25. Ank N, West H, Bartholdy C, Eriksson K, Thomsen AR, Paludan SR. Lambda interferon (IFN-lambda), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo. *J Virol*. 2006;80(9):4501–4509.

26. Ank N, Iversen MB, Bartholdy C, et al. An important role for type III interferon (IFN-lambda/IL-28) in TLR-induced antiviral activity. *J Immunol*. 2008;180(4):2474–2485.

27. Ank N, Paludan SR. Type III IFNs: new layers of complexity in innate antiviral immunity. *Biofactors*. 2009;35(1):82–87.

28. Liapakis A, Jacobson I. Pharmacogenetics of hepatitis C therapy. *Pharmacogenomics*. 2010;11(2):135–139.

29. Marcello T, Grakoui A, Barba-Spaeth G, et al. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology*. 2006;131:1887–1898.

30. Chen L, Borozan I, Feld J, et al. Hepatic gene expression discriminates between responders and nonresponders in treatment of chronic hepatitis C viral infection. *Gastroenterology*. 2005;128(5):1437–1444.

31. Chen L, Sun J, Meng L, et al. ISG15, a ubiquitin-like interferon stimulated gene, promotes Hepatitis C Virus production in vitro: Implications for chronic infection and response to treatment. *J Gen Virol*. 2010;91:382–388.

32. Sarasin-Filipowicz M, Oakeley EJ, Duong FH, et al. Interferon signaling and treatment outcome in chronic hepatitis C. *Proc Natl Acad Sci U S A*. 2008;105(19):7034–7039.

33. Jordan WJ, Eskdale J, Srinivas S, et al. Human interferon lambda-1 (IFN-lambda1/IL-29) modulates the Th1/Th2 response. *Genes Immun*. 2007;8:254–261.

34. Khakoo SI, Thio CL, Martin MP, et al. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science*. 2004; 305:872–874.

35. Gad HH, Dellgren C, Hamming OJ, Vends S, Paludan SR, Hartmann R. Interferon-lambda is functionally an interferon but structurally related to the interleukin-10 family. *J Biol Chem*. 2009;284:20869–20875.

36. Doyle SE, Schreckhise H, Khuu-Duong K, et al. Interleukin-29 uses a type 1 interferon- like program to promote antiviral responses in human hepatocytes. *Hepatology*. 2006;44:896–906.

37. Samarajiwa SA, Forster S, Auchettl K, Hertzog PJ. INTERFEROME: the database of interferon regulated genes. *Nucleic Acids Res*. 2009; 37(Database issue):D852–D857.

38. Pagliaccetti NE, Eduardo R, Kleinstein SH, Mu XJ, Bandi P, Robek MD. Interleukin-29 functions cooperatively with interferon to induce antiviral gene expression and inhibit hepatitis C virus replication. *J Biol Chem*. 2008;283(44):30079–30089.

39. Welzel TM, Morgan TR, Bonkovsky HL, et al. Variants in interferon-alpha pathway genes and response to pegylated interferon-Alpha2a plus ribavirin for treatment of chronic hepatitis C virus infection in the hepatitis C antiviral long-term treatment against cirrhosis trial. *Hepatology*. 2009;49(6):1847–1858.

40. Lin E, Hwang Y, Chen EY. Gene-gene and gene-environment interactions in interferon therapy for chronic hepatitis C. *Pharmacogenomics*. 2007;8(10):1327–1335.

41. Hijikata M, Ohta Y, Mishiro S. Identification of a single nucleotide polymorphism in the MxA gene promoter (G/T at nt-88) correlated with the response of hepatitis C patients to interferon. *Intervirology*. 2000;43(2):124–127.

42. Yee LJ, Tang J, Gibson AW, Kimberly R, Van Leeuwen DJ, Kaslow RA. Interleukin 10 polymorphisms as predictors of sustained response in antiviral therapy for chronic hepatitis C infection. *Hepatology*. 2001; 33(3):708–712.

43. Sugimoto Y, Kuzushita N, Takehara T, et al. A single nucleotide polymorphism of the low molecular mass polypeptide 7 gene influences the interferon response in patients with chronic hepatitis C. *J Viral Hepat*. 2002;9(5):377–384.

44. Kung SY, Hwang JN. Neural networks for intelligent multimedia processing. *Proceedings of the IEEE*. 1998;86:1244–1272.

45. Erb RJ. Introduction to backpropagation neural network computation. *Pharm Res*. 1993;10:165–170.

46. Serretti A, Smerald E. Neural network analysis in pharmacogenetics of mood disorders. *BMC Medical Genetics*. 2004;5:27.

47. Lin E, Chen PS, Lee IH, et al. Modeling short-term antidepressant responsiveness with artificial neural networks. *Open Access Bioinformatics*. In press 2010.

48. Ke WS, Hwang Y, Lin E. Pharmacogenomics of drug efficacy in the interferon treatment of chronic hepatitis C using classification algorithms. *Advances and Applications in Bioinformatics and Chemistry*. 2010 In press.

49. Lin E, Hwang Y. A support vector machine approach to assess drug efficacy of interferon-alpha and ribavirin combination therapy. *Mol Diagn Ther*. 2008;12(4):219–223.

50. Bishop CM. *Neural Networks for Pattern Recognition*. Oxford, UK: Clarendon Press; 1995.

51. Witten IH, Frank E. *Data mining: practical machine learning tools and techniques*. San Francisco, CA: Morgan Kaufmann Publishers; 2005.

52. Rumelhart DE, Hinton GE, William RJ. Learning internal representation by error propagation. In: *Parallel distributed processing: explorations in the micro-structure of cognition, vol.1: Foundations*. Cambridge, MA: MIT Press; 1996.

53. Huang LC, Hsu SY, Lin E. A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data. *J Transl Med*. 2009;7(1):81.

54. Kohavi R, John GH. Wrappers for feature subset selection. *Artificial Intelligence*. 1997;97:273–324.

55. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett*. 2006;27:861–874.

56. Hewett R, Kijsanayothin P. Tumor classification ranking from microarray data. *BMC Genomics*. 2008;9 Suppl 2:S21.

57. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23:2507–2517.

58. Sargent DJ. Comparison of artificial neural networks with other statistical approaches. *Cancer*. 2001;91(8):1636–1642.

59. Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Medical Informatics and Decision Making*. 2005;5:3.

60. Lin E, Hwang Y, Liang KH, Chen EY. Pattern-recognition techniques with haplotype analysis in pharmacogenomics. *Pharmacogenomics*. 2007;8(1):75–83.

61. White H. Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*. 1990;3:535–549.

62. Sontag ED. Feedback stabilization using two-hidden-layers nets. *IEEE Trans Neural Networks*. 1992;3(6):981–990.

63. Lin CT, Lee G. *Neural fuzzy systems*. Upper Saddle River, NJ: Prentice-Hall; 1996.