Liver Gene Expression Snap Shot

Sites Participating: All sites and all patients in the HALT-C Trial

Co-Chairs: Timothy Morgan, MD (University of California, Irvine), Herb Bonkovsky, MD (University of Connecticut)

Co-Investigators: Mary Carrington, PhD; Raymond Chung, MD; Michael Dean, PhD; Abhijit Dasgupta, PhD; Jules Dienstag, MD; Thomas O'Brien, MD, MPH; Sholom Wacholder, PhD; Denise Whitby, PhD; Chiuchin Yuan, PhD; Mingdong Zhang, MD, PhD

Study Name: Analysis of Gene Expression in HCV-infected Liver: Identification of Genes Associated with Responsiveness to Pegylated Interferon and Ribavirin Therapy

Separate Consent Form: No

Withdrawal Form: Yes (Form #9)

Eligible Patients: All Lead-in patients who provide consent for genetic testing. Twenty patients who are sustained responders and have frozen liver tissue will be compared to 20 matched controls.

Visit Schedule:
Frozen liver samples are sent to the repository if sufficient tissue is obtained during the screening biopsy.

Genetic Consent:
This consent is part of the main consent form for the HALT-C trial. Genetic consent is recorded on Form #4 and may be changed by the patient at any time, complete Form #9 Genetic Status Change. The current status of genetic consent is checked before DNA is prepared or samples are tested.

Analysis of Gene Expression in HCV-infected Liver: Identification of Genes Associated with Responsiveness to Pegylated Interferon and Ribavirin Therapy

Co-Investigators, HALT-C: Timothy Morgan, MD; Herb Bonkovsky, MD; Jules Dienstag, MD; Raymond Chung, MD

Co-Investigators, NCI Division of Cancer Epidemiology and Genetics: Mingdong Zhang, MD, PhD: Abhijit Dasgupta, PhD; Thomas O'Brien, MD, MPH

Co-Investigators, NCI Viral Epidemiology Laboratory SAIC: Denise Whitby, PhD; Chiuchin Yuan, PhD

Consultants, NCI Biostatistics Branch, Division of Cancer Epidemiology and Genetics: Sholom Wacholder, PhD

Consultants, NCI Laboratory of Genomic Diversity, Center for Cancer Research: Michael Dean, PhD; Mary Carrington, PhD

SUMMARY

Chronic hepatitis C virus (HCV) infection is a major cause of chronic liver disease and hepatocellular carcinoma. Therapy with pegylated interferon and ribavirin can resolve the infection, but many patients do not respond to this regimen. Preliminary findings suggest that higher serum levels of Th1-like cytokines are associated with poorer response to therapy. Host genetic factors may contribute to responsiveness to interferon treatment and recent technologic advances allow gene expression profiling of hepatic tissues. This study will test the hypothesis that expression of Th1-like cytokine genes is associated with non-response to pegylated interferon and ribavirin therapy. We also propose to explore the entire human genome to identify other hepatic gene transcripts that may be associated with responsiveness to therapy.

Subjects will be enrolled in the **H**epatitis C **A**ntiviral **L**ong-term **T**reatment against **C**irrhosis (HALT-C) trial. These patients have chronic HCV infection and advanced fibrosis. During the lead-in phase of HALT-C subjects receive pegylated interferon and ribavirin. Study subjects will consist of 20 patients who responded to this therapy and 20 virologic non-responders. We will use the Affymetrix GeneChip to compare hepatic gene expression in biopsy samples taken before the initiation of therapy. Cluster analyses will be performed to determine if responders can be separated from non-responders on the basis of expression of cytokine genes. We will also compare responders and non-responders with regard to the prevalence of cytokine gene expression and cytokine gene expression levels. We will use a similar strategy to explore the human genome for other genes that may be associated with treatment response. Identification of genetic determinants of response to anti-HCV therapy may lead to novel therapies and improve strategies for existing treatments. Improved treatment of HCV infection will help prevent the development of chronic liver disease and hepatocellular carcinoma.

BACKGROUND

Chronic infection with hepatitis C virus (HCV) affects about 3.9 million persons in the United States (Alter et al., 1999). Chronic HCV infection is a major cause of cirrhosis, end-stage liver diseases and hepatocellular carcinoma. An increase in the number of cases of hepatocellular carcinoma has occurred in the United States over the past two decades and HCV may account for most of the increase (El-Serag and Mason, 1999; El-Serag and Mason, 2000). Effective treatment of chronic hepatitis C can prevent the development of hepatocellular carcinoma. Interferon, either alone or in combination with other drugs, is the primary treatment available for chronic hepatitis C. However, less than half of the patients with chronic HCV infection respond to interferon therapy (Poynard et al., 1998; McHutchison et al., 1998). Pegylated interferon (interferon-alpha that has been conjugated with polyethylene glycol) has a significantly longer concentration peak than standard interferon-alpha (Glue et al., 2000). In patients with chronic hepatitis C, subcutaneous injection of pegylated interferon alpha once a week has been associated with a higher proportion of sustained virological response than administration of standard interferon-alpha three times a week, but with a similar safety profile (Heathcote et al., 2000; Zeuzem et al., 2000; Reddy et al., 2001). Early results from retreatment of previous non-responders to interferon or interferon and ribavirin therapy with pegylated interferon and ribavirin in HALT-C trial (see below) showed a 43% of virologic response rate (Shiffman, 2001).

Both viral and host cellular factors are thought to be involved in responsiveness to interferon treatment. Viral factors, such as serum HCV RNA levels, HCV subtype, diversity of the hypervariable region, and mutation of non-structural gene NS5A, and host factors, such as age, gender, duration of infection, alcohol intake, hepatic iron stores, platelet count, and histological staging of the liver disease have been proposed as predictors of response to interferon therapy (Shiratori and Omata, 2000; Polyak and Gerotto, 2000). The effect of host immune response, especially circulating immunoregulatory cytokines, in determining a patient's response to interferon therapy has also been studied (Saito et al., 2000; Bozkaya ett al., 2000; Sobue et al., 2001; Yee et al., 2001). Results from these studies are preliminary and somewhat inconsistent. It is postulated that the balance between Th1-like and Th2-like cytokines (Table 1) may be important in achieving therapeutic results. Th1-like cytokines, secreted by Th1 (inflammatory) T cells and other leukocytes, mediate inflammatory and cellular immune response. Th2-like cytokines, secreted by Th2 (helper) T cells and other leukocytes, modulate humoral immune response. Recently, Polyak et al. showed that elevated serum levels of IL-8 are associated with resistance to interferon therapy (Polyak et al., 2001). Neuman et al. found that individuals with sustained response to interferon therapy had lower TNF-α levels at baseline (Neuman et al., 2001). Serum IL-2R levels are also proposed as a serological marker of outcome following interferon treatment: its levels are more than 3-fold higher in non-responders than in responders (Grungrriff et al., 1999).

Table 1. List of Th1-like and Th2-like cytokines

Type	Cytokines
Th1-like	Interleukin (IL)-1, IL-2, IL-8, IL-12, TNF-α, TGF-β, IFN-γ
Th2-like	IL-4, IL-5, IL-6, IL-10, IL-13

Liver tissue samples available from the recently initiated **H**epatitis C **A**ntiviral **L**ong-term **T**reatment against **C**irrhosis (HALT-C) trial may make it possible to identify host cellular genes that are associated with responsiveness to pegylated interferon and ribavirin therapy. The HALT-C trial is a randomized, controlled trial to determine if long-term therapy with pegylated interferon can reduce the risk of histological progression to cirrhosis, liver decompensation, and hepatocellular carcinoma in patients with chronic hepatitis C and advanced fibrosis or cirrhosis (Figure 1). Enrollment in the study is limited to patients who have failed to respond to previous interferon therapy. Subjects with conditions potentially affecting response to the therapy, such as any other co-existent liver disease, co-infection with HIV, a platelet count < 75,000, a history

of alcohol abuse within the past year, and poorly controlled diabetes mellitus, are excluded from the study. In the lead in phase of the trial, all patients will receive 24 weeks of treatment with pegylated interferon and ribavirin. Those who fail to fully suppress HCV on this regimen will be randomized to receive either 42 months of pegylated interferon or no therapy. Liver biopsy will be performed on study subjects before the initiation of the therapy, and at 24 and 48 months after the start of the Lead-in Phase.

The recently developed Affymetrix GeneChip technology allows expression analysis of approximately 31,000 of the best-characterized genes in human cells or tissues. This set of two chips, to be released next spring, has extensive coverage of the human genome. It is based on Unigene 133 and cross referenced to GenBank, dbEST, RefSeq, Washington University EST Trace collection, and University of California Santa Cruz "golden-path" database. We propose to use the GeneChip technology to compare gene expression profiles in liver tissues from subjects who respond to pegylated interferon and ribavirin therapy and those who do not.

PROPOSED RESEARCH

Aims.

(1) To test the hypothesis that expression of Th1-like cytokine genes is associated with non-response to pegylated interferon and ribavirin therapy.
(2) To identify other host genes that are associated with responsiveness to pegylated interferon and ribavirin therapy.

Study design.

Subjects. Subjects who enroll in the HALT-C trial, complete 24 weeks of treatment with pegylated interferon and ribavirin therapy, and have an adequate frozen liver biopsy specimen available from a baseline study visit (i.e., prior to the onset of therapy during the lead in phase of the trial) will be eligible for this study. Liver biopsy samples, taken before the pegylated interferon and ribavirin therapy, from 20 sustained virologic responders and 20 non-responders, will be compared for hepatic gene expression profiles and expression levels. It would be more efficient to select subjects from a group as homogeneous as possible. Based on data from early enrollment, Caucasian men under 60 are likely to be the dominant group among study subjects. The responders and non-responders will be matched on the degree of liver fibrosis before the onset of the therapy. To eliminate the effect of the genotypes of HCV infected, we will limit the study to subjects infected with genotype 1 only.

Tissue specimens. Frozen liver tissue samples of approximately 10 mm (~4 mg), will be required for this study. A phenol-based isolation kit (Ambion, Austin, TX) will be used to extract total RNA from liver tissues. With this kit, the typical yield of total RNA from 4 mg of liver tissue is 8-12 μg.

Laboratory methods. Affymetrix GeneChips are oligonucleotide arrays that can be used to obtain the absolute expression level of genes in a sample (in contrast to cDNA microarrays which measure and compare differential gene expression between two samples). The computer controlled on-chip synthesis method used to create the oligonucleotide arrays typically results in a lower random error rate than is found for spotted cDNA array chips (Lipshutz et al. 1999).

We will follow the manufacturer's protocols. Briefly, total RNA from frozen hepatic tissue from each study subject will be extracted, purified, and quantified. Five μg of total RNA from each subject will be used to synthesize double-stranded cDNA with T7-$(dT)_{24}$ primer. Double-stranded cDNA will be purified and used in an *in vitro* transcription reaction to synthesize biotin-labeled cRNA. The labeled cRNA probe will then be purified, quantified, fragmented, and hybridized to the GeneChips (Affymetrix U95 set). Pre-mixed hybridization control transcripts

(*bioB*, *bioC*, *bioD*, and *Cre*) provided in the GeneChip Eukaryotic Hybridizatio n Control Kit (Affymetrix) will be used as positive controls. The hybridized chips will be washed, stained, and scanned. The signal is read using a fluorimager device. Several hybridization controls are used to generate normalized expression levels that allow comparison of RNA levels among samples.

All chips used in the experiments will be ordered from the same production lot to minimize variability in the production of GeneChips. Additionally, we will replicate a few hybridizations to make sure that measurements of gene expression levels are reproducible. We will not extract RNA from HALT-C liver specimens until we have reproducibly isolated RNA, and successfully performed the Affymetrix GeneChip assay, on frozen liver specimens from other sources.

Data analysis.

General calculation. We will use GeneChip Analysis Suite software (Affymetrix, Santa Clara, CA) to determine (1) if a gene transcript is expressed (the "Presence" or "Absence") and (2) the expression level of the transcript. On a GeneChip, each gene transcript is represented by a "probe set" made up of 16-20 pairs of oligonucleotide "probes." Each pair consists of a 25 base probe that is perfectly complementary to a specific gene sequence ("Perfect Match" or PM) and a probe that differs from the PM by a single base at the central (13^{th}) position ("Mismatch" or MM). MM probes serve as a control for cross-hybridization by transcripts from other genes. The determination of the "Presence" or "Absence" of each gene transcript is based on a decision matrix that considers the relative intensity among pairs of PM and MM probes, and the background intensity of the entire array. The expression level of each gene transcript is determined by calculating the *Average Difference* between the PM and MM pairs that make up the probe set.

Hypothesis testing. To test the hypothesis that expression of Th1-like cytokine genes is associated with non-response to pegylated interferon and ribavirin therapy, we will first perform cluster analyses on all expression profiles of Th1-like and Th2 like cytokine genes to test if responders and non-responders are clustered as separate groups, and to identify cytokine genes that are differentially expressed between the responders and non-responders. Hierarchical clustering is a bottom-up procedure that places similar data close together, organizing them into a mock-phylogenetic tree or dendrogram. The distance between two branches is a measure of the correlation between the groups of data points represented by the two branches. Such a procedure can be used either on the gene level to cluster genes with similar patterns across samples or to cluster samples based on their overall gene expression. Hierarchical clustering procedures contained in the GeneSpring™ software (Silicon Genetics, Redwood City, CA) will be used to see whether the samples from responders and non-responders cluster into separate groups, providing an indication of genetic differences between the groups. Such clustering procedures will also be used on the genes to identify possible groupings for genes with similar expression patterns. Hierarchical clustering produces a greater number of smaller clusters, which aids in class discovery (Tibshirani et al, 1999). We will also use a clustering procedure called "gene shaving" (Hastie et al, 2000) which finds clusters of genes that exhibit similar expression patterns within groups and different expression patterns between groups. Such a procedure can help in identifying genes that are differentially expressed between the responders and non-responders.

In addition, we will compare the prevalence of "Presence" of each cytokine gene between responders and non-responders by Chi-square test. A gene with a difference of 0.5 in prevalence between responders and non-responders is considered differentially expressed. The pattern of differentially expressed Th1-like and Th2-like cytokine genes will be analyzed to determine if imbalance of Th1-like and Th2-like cytokine gene expression is associated with responsiveness to the therapy.

We will also compare the expression level of each cytokine gene between responders and non-responders by Student t test. Mathematical transformation of data will be performed if necessary. This analysis will test if expression levels of certain cytokine genes are associated with responsiveness to the therapy.

A type 1 pattern of cytokine gene expression in a group (responders or non-responders) is defined as an expression of IL-12, IFN-γ, and other Th1-like cytokine genes, coupled with the absence of Th2-like cytokine genes. A type 2 pattern is defined as an expression of IL-4, IL-10, and other Th2-like cytokine genes, coupled with the absence of the Th1-like cytokine genes. If both Th1-like and Th2-like cytokine genes are expressed in a group, the percentage of each type of cytokine genes present will be compared. Expression levels (fold changes) of each type of cytokine genes present will also be compared. The type with a higher percentage of cytokine genes present and higher average expression levels of genes present is considered the dominant type.

Data exploration. We will first perform cluster analyses on all expression profiles to test if responders and non-responders are clustered as separate groups, and to identify genes that are differentially expressed between the responders and non-responders.

The same cluster analyses will be repeated on subsets of genes sifted according to their presence or expression levels. First, we will limit the analyses to genes that are present. Then we will sift genes based on their differences in expression levels. We will start analyses with genes that show a 10-fold change in expression levels between responders and non-responders, and move to genes with a 5-fold and 3-fold change in expression levels. By pulling out genes with specified expression levels and limiting analysis to genes that are present, we can use more stringent criteria to identify genes that are differentially expressed between responders and non-responders.

We will then compare the prevalence of "Presence" of each candidate gene identified by cluster analyses between responders and non-responders by Chi-square test. A gene with a difference of 0.5 in prevalence between responders and non-responders is considered differentially expressed (see statistical power below).

We will also compare the expression level of each candidate gene between responders and non-responders by student t test for paired samples. Mathematical transformation of data will be performed if necessary. This analysis will test if expression levels of candidate genes are associated with responsiveness to the therapy.

Statistical power.
For each gene we will have at least 80% power to detect a difference of 0.45 and at least 90% power to detect a difference of 0.50 in the prevalence of a particular gene among responders to the prevalence of that gene among non-responders, under a Type I error of 0.05.

Multiple comparisons are a concern when 31,000 genes are studied. By identifying candidate genes through cluster analyses and by limiting analyses to genes that are present and to genes with a stringent fold change in expression levels, we hope to minimize statistically significant candidate genes, and therefore to minimize the effect of multiple comparisons to some extent. Besides, the nature of the proposed study is exploratory and the purpose of the study is to generate more specific hypotheses. Findings from this study will provide leads for future confirmatory investigations.

CONTRIBUTION OF COLLABORATORS

Dr. Zhang will design the study, perform and oversee the laboratory work, and analyze the data.

Dr. Dasgupta will develop novel approaches to analyze microarray data, help design the study, and oversee the data analyses.

Dr. O'Brien will supervise every aspect of the study.

Drs. Bonkovsky, Chung, Dienstag, and Morgan are co-investigators of the HALT-C study, and will coordinate biopsy tissue collections and oversee human subject protection.

Ms. Yuan will help perform the GeneChip assays; Dr. Whitby will oversee this work.

Drs. Dean and Carrington will help interpret the results.

Dr. Wacholder will be a consultant on data analysis.

REFERENCES

Alter MJ, Kruszon-Moran D, Nainan OV, McQuillan GM, Gao F, Moyer LA, Kaslow RA, Margolis HS. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. N Engl J Med 1999 Aug 19;341(8):556-62.

Bozkaya H, Bozdayi AM, Aslan N, Turkay C, Sarioglu M, Cetinkaya H, Akdogan M, Cinar K, Erden E, Kose K, Senturk H, Akkiz H, Karayalcin S, Yurdaydin C, Uzunalimoglu O. Circulating IL-2 and IL-10 in chronic active hepatitis C with respect to the response to IFN treatment. Infection. 2000 Sep;28(5):309-13.

El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999 Mar 11;340(10):745-50.

El-Serag HB, Mason AC. Risk factors for the rising rates of primary liver cancer in the United States. Arch Intern Med 2000 Nov 27;160(21):3227-30.

Glue P, Fang JW, Rouzier-Panis R, Raffanel C, Sabo R, Gupta SK, Salfi M, Jacobs S. Pegylated interferon-alpha2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. Hepatitis C Intervention Therapy Group. Clin Pharmacol Ther. 2000 Nov;68(5):556-67.

Grungreiff K, Reinhold D, Ansorge S. Serum concentrations of sIL-2R, IL-6, TGF-beta1, neopterin, and zinc in chronic hepatitis C patients treated with interferon-alpha. Cytokine. 1999 Dec;11(12):1076-80.

Hastie T, Tibshirani R, Eisen MB, Alizadeh A, Levy R, Staudt L, Chan WC, Botstein D, Brown P. 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. Genome Biology. 2000, 1(2):research0003.1-0003.21

Heathcote E. J., Shiffman M. L., Cooksley W. G. E., Dusheiko G. M., Lee S. S., Balart L., Reindollar R., Reddy R. K., Wright T. L., Lin A., Hoffman J., De Pamphilis J. Peginterferon Alfa-2a in Patients with Chronic Hepatitis C and Cirrhosis. N Engl J Med 2000; 343:1673-1680, Dec 7, 2000.

Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. Nat Genet. 1999 Jan;21(1 Suppl):20-4.

Neuman MG, Benhamou JP, Malkiewicz IM, Akremi R, Shear NH, Asselah T, Ibrahim A, Boyer N, Martinot-Peignoux M, Jacobson-Brown P, Katz GG, Le Breton V, Le Guludec G, Suneja A, Marcellin P.Cytokines as predictors for sustained response and as markers for immunomodulation in patients with chronic hepatitis C.
Clin Biochem. 2001 May;34(3):173-82.

McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, Goodman ZD, Ling MH, Cort S, Albrecht JK. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. N Engl J Med. 1998 Nov 19; 339(21):1485-92.

Polyak SJ, Gerotto M. The molecular basis for responsiveness to anti-viral therapy in hepatitis C. Forum (Genova). 2000 Jan-Mar;10(1):46-58. Review.

Polyak SJ, Khabar KS, Rezeiq M, Gretch DR. Elevated levels of interleukin-8 in serum are associated with hepatitis C virus infection and resistance to interferon therapy.
J Virol. 2001 Jul;75(13):6209-11.

Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, Bain V, Heathcote J, Zeuzem S, Trepo C, Albrecht J. Randomized trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). Lancet. 1998 Oct 31;352(9138):1426-32.

Reddy KR, Wright TL, Pockros PJ, Shiffman M, Everson G, Reindollar R, Fried MW, Purdum PP 3rd, Jensen D, Smith C, Lee WM, Boyer TD, Lin A, Pedder S, DePamphilis J. Efficacy and safety of pegylated (40-kd) interferon alpha-2a compared with interferon alpha-2a in noncirrhotic patients with chronic hepatitis C. Hepatology. 2001 Feb;33(2):433-8.

Saito H, Ebinuma H, Satoh I, Miyaguchi S, Tada S, Iwabuchi N, Kumagai N, Tsuchimoto K, Morizane T, Ishii H. Immunological and virological predictors of outcome during interferon-alpha therapy of chronic hepatitis C. J Viral Hepat. 2000 Jan;7(1):64-74.

Shiffman ML. Retreatment of interferon and interferon-ribavirin non-responders with peginterferon-alpha-2a and ribavirin: initial results from the lead-in phase of the HALT-C trial. Hepatology 2001 October; 34(4):243A.

Shiratori Y, Omata M. Predictors of the efficacy of interferon therapy for patients with chronic hepatitis C before and during therapy: how does this modify the treatment course? J Gastroenterol Hepatol. 2000 May;15 Suppl:E141-51.

Sobue S, Nomura T, Ishikawa T, Ito S, Saso K, Ohara H, Joh T, Itoh M, Kakumu S. Th1/Th2 cytokine profiles and their relationship to clinical features in patients with chronic hepatitis C virus infection. J Gastroenterol. 2001 Aug;36(8):544-51.

Tibshirani R, Hastie T, Eisen M, Ross D, Botstein D, Brown P. Clustering methods for the analysis of DNA microarray data. 1999. Technical report, Department of Statistics, Stanford University.

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 Mar;33(3):708-12.

Zeuzem S., Feinman S. V., Rasenack J., Heathcote E. J., Lai M.-Y., Gane E., O'Grady J., Reichen J., Diago M., Lin A., Hoffman J., Brunda M. J. Peginterferon Alfa-2a in Patients with Chronic Hepatitis C. N Engl J Med 2000; 343:1666-1672, Dec 7, 2000.