HALT-C Ancillary Study PROPOSAL

Part I

Proposal Name: Validation of Cirrhosis Risk Score in Patients with Chronic Hepatitis C

Proposal PI: Hongjin Huang, PhD, Celera

HALT-C PI: Anna S. Lok, MD

Funding Agency and Review Body (e.g., NIDDK; my university/GAC): Celera Diagnostics

I agree to follow HALT-C Policies and Procedures when conducting this study. I acknowledge that the data obtained from this study will belong to the NIH and will be placed in the HALT-C database for use by other investigators. I understand that I cannot begin experiments using HALT-C specimens/data until I receive approval from the HALT-C Steering Committee and funding from the Scientific Review Body for my proposal. I also understand that the data analysis for this proposal will be performed by NERI (unless otherwise approved by the HALT-C study) and that protocols approved by the HALT-C Steering Committee will be placed on the HALT-C Restricted Website.

_____ Proposal Principal Investigator	_____ Date
_____ HALT-C Principal Investigator Date	_____ Date

Protocol Part II (4 page limit, single space)

1. Aims/hypotheses

Hypothesis

- Genetic factors play a role in fibrosis progression and adverse clinical outcomes (hepatic decompensation and hepatocellular carcinoma [HCC]) in patients with hepatitis C

Primary aims

- To validate the Caucasian cirrhosis risk score (CRS) in predicting the risk of cirrhosis in HALT-C non-Hispanic Caucasian patients and to evaluate the accuracy of the Caucasian CRS in predicting the risk of progression from advanced fibrosis to cirrhosis and the risk of development of cirrhosis complications.
- To evaluate the Caucasian CRS in predicting the risk of cirrhosis and cirrhosis complications in HALT-C African American patients.

Secondary aims

- To validate the African American CRS (when available) in HALT-C African American patients.
- To evaluate the combined utility of CRS and non-invasive indices of hepatic fibrosis such as AST-platelet ratio index (APRI) and the HALT-C cirrhosis prediction index.

2. Background/rationale

- The fibrosis progression rate is highly variable among patients with chronic hepatitis C. Overall it is estimated that only about 20% of patients with chronic hepatitis C will progress to cirrhosis after 20 years of infection[1]. Previously identified clinical risk factors for fibrosis progression include male gender, excessive alcohol use, older age at the time of infection and the presence of steatosis on liver biopsy[2]. However, many patients with these characteristics have mild disease and some patients without these factors have cirrhosis[3].
- Besides clinical factors, recent studies suggested that host genetic factors, such as single nucleotide polymorphisms (SNPs) could play an important role in determining the risk for fibrosis progression[4,5].
- In the past 5 years, Celera Diagnostics has worked with 5 major academic centers in the US and enrolled 1,468 patients (984 Caucasians, 293 African Americans and 191 other races) with chronic hepatitis C (CHC). From a functional genomic scan of 25,000 SNPs, 361 SNPs were selected for the signature building based on their significant association with fibrosis risk in two large cohorts[5]. Using a 'machine learning' approach, a signature consisting of 7 markers most predictive of cirrhosis risk in Caucasian patients was developed, and Cirrhosis-Risk-Score (CRS) was calculated to determine the risk for each patient.
- The area-under-the-ROC-curves (AUROC) of the CRS in Caucasian patients was 0.75 in the Training cohort. In the Validation cohort, AUROC was only 0.53 for clinical risk factors, increased to 0.73 for CRS, and to 0.76 when CRS was combined with clinical risk factors. A CRS low-cutoff of <0.50 to identify low-risk patients would misclassify only 10.3% of high-risk patients, while a high-cutoff of >0.73 to identify high-risk patients would misclassify 22.3% of low-risk patients (Table 1).

Table 1: Predictive Values of CRS

5A. Training Cohort

CRS	All Patients	Low-Risk (Stage 0) (N=157)	High-Risk (Stage 3-4) (N=263)	NPV	PPV	Sensitivity	Specificity	Misclassifying Rate[a]
< 0.10	4 (1.0%)	4	0	100.0%		100.0%	2.5%	0.0%
< 0.20	18 (4.3%)	16	2	88.9%		99.2%	10.2%	0.8%
< 0.30	30 (7.1%)	27	3	90.0%		98.9%	17.2%	1.1%
< 0.40	55 (13.1%)	46	9	83.6%		96.6%	29.3%	3.4%
< 0.50[b]	98 (23.3%)	71	27	72.4%		89.7%	45.2%	10.3%
> 0.70[b]	194 (46.2%)	35	158		81.4%	60.1%	77.7%	22.3%
> 0.75	155 (36.9%)	28	127		81.9%	48.3%	82.2%	17.8%
> 0.80	67 (16.0%)	11	56		83.6%	21.3%	93.0%	7.0%
> 0.85	55 (13.1%)	9	46		83.6%	17.5%	94.3%	5.7%
> 0.90	10 (2.4%)	0	10		100.0%	3.8%	100.0%	0.0%

5B. Validation Cohort

CRS	All Patients	Low-Risk (Stage 0) (N=14)	High-Risk (Stage 3-4) (N=140)	NPV	PPV	Sensitivity	Specificity	Misclassifying Rate[a]
< 0.10	1 (0.6%)	1	0	100.0%		100.0%	7.1%	0.0%
< 0.20	5 (3.2%)	3	2	60.0%		98.6%	21.4%	1.4%
< 0.30	14 (9.1%)	6	8	42.9%		94.3%	42.9%	5.7%
< 0.40	16 (10.4%)	6	10	37.5%		92.9%	42.9%	7.1%
< 0.50[b]	23 (14.9%)	6	17	26.1%		87.9%	42.9%	12.1%
> 0.70[b]	78 (50.6%)	3	75		96.2%	53.6%	78.6%	21.4%
> 0.75	57 (37.0%)	3	54		94.7%	38.6%	78.6%	21.4%
> 0.80	31 (20.1%)	1	30		96.8%	21.4%	92.9%	7.1%
> 0.85	26 (16.9%)	1	25		96.2%	17.9%	92.9%	7.1%
> 0.90	4 (2.6%)	0	4		100.0%	2.9%	100.0%	0.0%

a. The low cutoff was chosen to minimize misclassification of patients at high risk of cirrhosis as being low risk

b. Using 0.50 and 0.70, 30.5% and 34.5% patients were not classifiable in Training and Validation cohort, respectively.

- Indices derived from routinely available laboratory results have been reported to correlate with advanced fibrosis or cirrhosis on biopsies. These indices include APRI (derived from AST and platelet)[6] and the HALT-C cirrhosis prediction index (derived from platelet, AST/ALT ratio and INR)[7].
- The CRS was developed with the aim to predict the risk of cirrhosis. To date, studies performed by Celera had been cross-sectional studies showing that the CRS correlated with histological cirrhosis with an AUROC or 0.73-0.75. To validate that the CRS can predict the risk of cirrhosis, Celera has embarked on a study involving multiple liver centers to recruit patients who had baseline liver biopsies at early disease stages (F0-F1) and who have undergone serial liver biopsies.
- The CRS did not use complications of cirrhosis (eg. hepatic decompensation and HCC) as endpoints for the study. It will be interesting to determine if the genes that predict cirrhosis also predict progression to cirrhosis complications.

3. Relations to aims of HALT-C study

The main aim of the HALT-C study is to determine if maintenance interferon therapy can delay disease progression in patients with hepatitis C and advanced fibrosis/cirrhosis, who failed to achieve sustained virologic response to antiviral therapy. Several ancillary studies were designed to examine noninvasive assessment of liver fibrosis, risk factors for cirrhosis as well as progression to decompensated cirrhosis and HCC. This proposal will complement those ancillary studies and will determine if the majority of the HALT-C patients will be predicted to develop cirrhosis based on genetic factors. The application of the Celera CRS to the subset of patients who progressed from bridging cirrhosis at entry to cirrhosis during the course of the HALT-C study will be of particular interest. The Celera CRS was built on data from Caucasian patients, the large number of African Americans with cirrhosis/advanced fibrosis in the HALT-C study provides an opportunity to evaluate the Caucasian CRS for African Americans with hepatitis C.

This proposal will also explore the utility of the Caucasian CRS in predicting decompensated cirrhosis and HCC. Additional discovery work including identification of other genes and formulation of new risk scores may be necessary for the prediction of decompensated cirrhosis and HCC. These new scores will need to be validated in external cohorts.

The current proposal will include exploration of the utility of the Caucasian CRS in predicting decompensated cirrhosis and HCC. A separate proposal will be developed in parallel to conduct additional discovery work to develop risk scores for predicting decompensated cirrhosis and HCC among patients with chronic hepatitis C. It is anticipated that the second proposal will entail more resources and a much longer period of time to complete, including possibly expanding PBMCs from patients who do not have adequate DNA in our repository. The goal is to be ready to launch the second proposal when work related to the first proposal is completed, assuming that the results of the first proposal are positive or encouraging.

4. Study design, experimental groups

Patients

- Non-Hispanic Caucasian and African American patients who participated in the HALT-C Study and who consented to genetic studies, patients of other racial/ethnic groups will not be included as their numbers are too small to generate meaningful data
- This study will include all patients who participated in the randomized phase of HALT-C as well as patients who were responders in the lead-in phase, and will include patients with bridging fibrosis as well as those with cirrhosis

Study design

- 250 ng genomic DNA from each eligible patient will be sent to Celera for genotyping

5. Methods, data usage

Required clinical data from HALT-C DCC

- A defined data set on patients who meet eligibility criteria listed in section 4 and in whom DNA samples are available will be created by NERI and sent to Celera for analysis. The dataset will include:
- (a) Demographics, duration of HCV infection, alcohol and smoking history, BMI, HCV genotype
- (b) Baseline information: histology (inflammation, fibrosis and steatosis score, length of biopsy and fragmentation), baseline labs (CBC, liver panel, INR, HCV RNA), baseline ultrasound finding (spleen size), (W24 or R00) EGD findings (presence or absence of varices/portal gastropathy)
- (c) M24 labs (CBC, liver panel, INR), ultrasound finding (spleen size) and histology (inflammation, fibrosis and steatosis score, length of biopsy and fragmentation)
- (d) M48 labs (CBC, liver panel, INR), ultrasound finding (spleen size), histology (inflammation, fibrosis and steatosis score, length of biopsy and fragmentation) and EGD findings (presence or absence of varices/portal gastropathy)
- (e) Occurrence of clinical outcomes
- DNA samples and dataset will be de-identified and coded using subject number only.

Genotyping

- CRS Multiplex Research Prototype: A single tube multiplex PCR-Oligonucleotide Ligation research assay (PCR-OLA) was established to genotype all 7 SNPs in CRS signature plus gender. Each of the eight distinct amplicons is genotyped by an OLA which involves the joining of an Allele-Specific Oligonucleotide containing a complementary sequence tag for a specific oligonucleotide-conjugated Luminex® bead with a Ligation-Specific Oligonucleotide containing a 3' biotin. The ligation products are then hybridized to derivatized Luminex® beads, labeled by the reporter molecule Streptavidin-Phycoerythrin, and analyzed on a Luminex® 100™ system. Genotypes are determined automatically by generic research software that compares the ratio of fluorescent signals between the two alleles.
- Allele Specific PCR: If additional SNPs needed to be genotyped, allele specific PCR will be used. For each SNP, two amplification reactions will be performed, one for each allele. Each PCR reaction contains one of the two allele-specific primers and a common reverse primer. All reactions will be amplified on Applied Biosystems PRISM 7900HT Sequence Detection instrument. For individual DNA, the genotypes will be automatically called by a custom-designed algorithm.

Data analyses

For all the following analyses, CRS value will be calculated using a Naïve Bayes algorithm developed by Celera (see details below). The performance of CRS in HALT-C patients will be evaluated by statistical parameters such as Area under the ROC Curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) etc.

- Stage 1A: Validate the current Caucasian CRS with 7 SNPs plus gender in Caucasian and African American patients from HALT-C trial in predicting
    - Cirrhosis on baseline biopsy
    - Progression to cirrhosis on M24 or M48 biopsy – for those who did not have cirrhosis on baseline biopsy

  Stage 1B: Explore the current Caucasian CRS with 7 SNPs plus gender in Caucasian and African American patients from HALT-C trial in predicting
    - Progression to cirrhosis complications – hepatic decompensation and/or HCC
- Stage 2: Validate the optimized Caucasian CRS (upon availability) in Caucasian and African American patients from HALT-C trial. Celera is performing fine density mapping of CRS signature, in which over 600 SNPs around the original 7 SNPs are being evaluated. These SNPs will be incorporated into the modeling process to obtain a more optimized Caucasian CRS signature. The optimized Caucasian CRS is expected to be available by June 2007.
- Stage 3: Validate the African American CRS (upon availability) in African American patients from HALT-C trial. Currently, Celera has DNA samples from 293 African American (AA) patients enrolled using the same criteria as Caucasian patients. Celera is in the process of identifying SNPs significant in AA population from the genome scan data. Although sample size may be inadequate, Celera will explore the possibility of building the AA CRS signature. The preliminary result is expected to be available in August 2007. If a predictive signature can be identified, its performance will be validated in AA patients from HALT-C trial.
- Stage 4: Evaluate the utility of combining the Celera CRS with age or duration of infection and a cirrhosis index based on laboratory test results at the time of evaluation in predicting cirrhosis.

Naïve Bayes Algorithm (optional)

The value of CRS based on a constellation of 7 SNPs was calculated using a Naïve Bayes formula13. Given outcomes C = {cirrhosis, no cirrhosis} and a set of seven predictive SNPs X = {X1, X2, … , X7}, the probability of a patient S = {X1 = x1, X2= x2, … , Xn= x7} having cirrhosis is computed as follows:

$$\mathbf{CRS} = \frac{0.626 * P(S \mid \text{cirrhosis})}{0.626 * P(S \mid \text{cirrhosis}) + 0.374 * P(S \mid \text{no Cirrhosis})}$$

The conditional probabilities P (S| cirrhosis) and P (S| no cirrhosis), were estimated assuming each SNP was independent of all other SNPs.

The probabilities, $P(X_i = x_i \mid cirrhosis)$, $P(X_i = x_i \mid no\ cirrhosis)$ are the class conditional probabilities for the ith marker Xi with value xi (each SNP can take the value of '1' or '0' based on the genotypes), given that the patient has cirrhosis or no cirrhosis respectively.

6. Anticipated results

Caucasian patients - We anticipate the vast majority of patients will be predicted to have cirrhosis according to the CRS, and the accuracy will be highest among those with cirrhosis in the baseline biopsy followed by those who progressed from bridging fibrosis on baseline biopsy to cirrhosis on M24 or M48 biopsies or developed clinical outcome by M48, and patients who had bridging fibrosis on baseline biopsy and did not progress to cirrhosis on M24 or M48 biopsies or develop clinical outcome by M48. We anticipate that the current CRS will be more accurate in predicting histological cirrhosis than hepatic decompensation or HCC. The risk score may need to be modified in order to effectively predict cirrhosis complications as well.

Deleted:

African American patients – Due to the genetic variations commonly seen in different races, it is hard to predict the performance of Caucasian CRS in African American patients. The results will be valuable to determine if additional studies are needed to generate a separate signature for African Americans. Data from African American HALT-C patients will also be used to validate the African American CRS when this is available.

Combination of CRS and age (duration of HCV infection) and a cirrhosis index based on laboratory results at the time of evaluation might further improve the accuracy of cirrhosis prediction.

7. Statistical support

Biostatisticians at Celera will perform the analysis, the final results will be verified by statisticians at NERI

8. HALT-C samples to be used in the study (complete Part III: Sample Requirements)

250 ng genomic DNA from each Caucasian and each African American patient in the HALT-C study, who consented to genetic studies and in whom DNA samples are available

9. Financial issues (e.g., cost for data analysis and obtaining samples from Repository)

Celera will cover the costs of identification, retrieval and shipment of samples from the Repository; preparation and transfer of the defined data set, statistical oversight, and coordination of the entire ancillary study.

10. References.

1. Thomas D and Seeff L 2005, Clin Liver Dis 9:383-98
2. Poynard T at al 1997, Lancet 349: 825-832
3. Marcellin P et al 2002, Hepatology 36:S47-56
4. Bataller R, Brenner DA. Liver fibrosis. J Clin Invest 2005;115:209-218.
5. Huang H, Shiffman ML, Cheung RC, Layden TJ, Friedman S, Abar OT, Yee L, et al. 2006, Gastroenterology 130:1679-1687.
6. Wai C et al 2003, Hepatology 38: 518-526
7. Lok A et al 2005; Hepatology 42: 282-292

Protocol Part III: Sample Requirements

Visit	Liver # patients, mm*	Blood # patients, ml	Other (describe) # pts, amount
Screen 1			Up to 670 samples of 250 ng genomic DNA from S00 or later visit.
Screen 2			
Baseline			
Lead in Week 4			
Week 8			
Week 12			
W16			
Week 20			
Week 24			
Randomized Month 9			
Month 12			
Month 15			
Month 18			
Month 21			
Month 24			
Month 27			
Month 30			
Month 33			
Month 36			
Month 39			
Month 42			
Month 45			
Month 48			
Post-treatment			