[Frontiers in Bioscience, Landmark, 21, 479-486, January 1, 2016]

A 6 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis B

Ming-Yi Xu[1], Ying Qu[1], Zheng-Hong Li[1], Fei Li[1], Chun-Yang Xiao[1], Lun-Gen Lu[1]

[1]Department of Gastroenterology, Shanghai First People's Hospital, Shanghai Jiao Tong University, Shanghai, 200080, China

TABLE OF CONTENTS

1. ABSTRACT

Clinical factors and liver biopsy cannot accurately predict the risk of developing cirrhosis in chronic hepatitis B (CHB).This study was to develop a predictive gene signature for cirrhosis in CHB patients. A total of 183 untreated CHB patients were enrolled. GeneChip, significant analysis of microarray (SAM) and prediction analysis of microarray (PAM) were used to select predictor genes (PGs) in liver tissues. The Cirrhosis Risk Score (CRS) was calculated based on 6 PG variables and the predictive value of CRS was evaluated. Firstly differentially expressed genes were filtered from a genome scan and SAM, and 87 significant genes were selected for the signature building. Secondly a signature consisting of 6 PGs (CD24, CXCL6, EHF, ITGBL1, LUM and SOX9) most predictive for cirrhosis risk in CHB patients was developed in the selection set (n=40) by use of PAM and PCR approach. Finally the CRS was calculated to estimate the risk of developing cirrhosis and then tested in validation cohort (n=143). The area under the ROC curves (AUROC) of the CRS was 0.944 and exceeded to 6 PGs and clinical factors. A low CRS cutoff of <6.43 to identify low-risk patients would misclassify only 8.16% of high-risk patients, while a high cutoff of >8.32 to identify high-risk patients would misclassify 0% of low-risk patients. So CRS is a better predictor than clinical factors in differentiating high-risk versus low-risk for cirrhosis and application of CRS in clinical practice could help to reduce the rate of liver biopsy in patients with CHB.

2. INTRODUCTION

Chronic hepatitis B (CHB) is the most common cause of cirrhosis and hepatocellular carcinoma (HCC), and the leading indication for liver transplantation in China. It has also been reported that about 400 million people are living with CHB worldwide, while 0.5 million people die from its related complications every year (1). The progression rate to cirrhosis varies widely among CHB patients. Treatment with antiviral therapy would be most cost-effective in those patients with evidence of progressive liver disease. Previously identified standard clinical risk factors for rapid progression include active viral replication (hepatitis B virus (HBV) DNA (>10^5-10^6 copies/ml)) or hepatitis B e antigen (HBeAg), alcohol use, co-infection with other viruses and use of immunosuppressive agents in CHB patients (2). However, due to individual variability, these factors are

not sufficiently accurate to identify which CHB patients with these factors will progress to cirrhosis (3).

We hypothesized that host genetic factors, such as hepatic gene expressional signature, could play a primary role in determining cirrhosis risk. We previously reported on the existence of a signature of gene expressions involved in liver fibrotic tissues of therapy-naive CHB patients, by use of microarray (4-5). This raised the possibility that this approach can be used to identify a specific condition, such as the risk of developing cirrhosis, in CHB patients. More importantly, a key question remained as to how these gene markers could be utilized in clinical settings. In the current study, we completed the confirmation of significant genes from a genomic scan, and selected genes for signature building. The aim of this study was to build and validate a gene signature that could be utilized in clinical practice to assess the risk for cirrhosis in CHB patients.

3. METHODS

3.1. Patients

Total 183 CHB patients were consecutively recruited from our hospital in 2008~2010. The enrollment criteria were based on established Chinese diagnostic guideline of 2005. Patients who had HBV infection with positive hepatitis B surface antigen (HBsAg) for at least 6 months were considered for inclusion in this study. In brief, all patients were adults (ages 18-60 years), HBsAg and HBV DNA positive, and had a baseline liver biopsy, untreated for HBV. Patients were excluded if they had any other co-existing chronic liver diseases or co-infection with human immunodeficiency virus (HIV) or hepatitis C virus (HCV), or consumed >30g of alcohol per day, or presence of HCC. The study was approved by the institutional review board of our centre and written informed consent was obtained from all patients.

3.2. Clinical risk factors and clinical endpoint

Medical records were reviewed regarding demographics, risk factors (and year) for HBV infection, virological features (HBeAg and HBV DNA) and liver histology were scored using the system described by Scheuer (6-7). The primary objective was to determine the risk of developing bridging fibrosis/cirrhosis (S3-4) in patients with CHB. To increase the signal-to-noise ratio in the modeling process, only those patients with histological status at the two extremes were included (8). CHB patients developing either S3-4 are considered as being 'high-risk' for developing cirrhosis. In contrast, patients with no-fibrosis (S0) are considered as being 'low-risk'.

3.3. Sample collection and pathological data

Hepatic tissue sections were processed with hematoxylin and eosin, Masson's trichrome, and reticular fiber scanning. A minimum length of at least 1.5 cm of the tissue obtained from liver biopsy and at least 6 portal tracts were required for diagnosis. Fibrotic stages and inflammation grades were determined using Scheuer's classification. All the sections were assessed by 3 pathologists in a blind and independent fashion.

3.4. Microarray hybridization and data analysis

Human U133 plus 2.0 array (Affymetrix, Santa Clara, CA) were performed in liver tissues of selection cohort (n=40). Data was normalized by applying the GC robust multi-array average algorithm (9) and the baseline was calculated using the geometric mean from the data of the controls. The raw data was analyzed and processed with GeneSping GX software (zcomSilicon Genetics, Redwood City, CA, USA).

3.5. Building cirrhosis risk score (CRS) signature

For selection of significant gene (SG) as candidate markers, significant analysis of microarray (SAM) (10) was included in selection cohort. The candidate markers were then analyzed by prediction of microarray analysis (PAM) (http://www-stat.stanford.edu/~tibs/PAM) (11) to generate a set of predictor genes (PGs). The selected PGs possessed the signature of cirrhosis risk. The PGs were further tested via qRT-PCR in liver tissues of selection cohort. The value of CRS based on 6 PG variables with fibrosis extent was assessed utilizing logistic multivariable linear regression. Given outcomes C= {cirrhosis, no cirrhosis} and F= {fibrosis, no fibrosis}, the probability of a patient having cirrhosis is computed. The value of Fibrosis Risk Score (FRS) was also assessed.

$$CRS1=0.273\times CD24+0.184\times CXCL6+0.209\times EHF+0.036\times SOX9+0.076\times ITGBL1$$

$$CRS2=0.036\times CXCL6+0.093\times EHF-0.111\times SOX9-0.008\times ITGBL1$$

$$FRS = -0.088\times CD24+0.294\times EHF+0.043\times LUM+0.081\times SOX9+0.109\times ITGBL1$$

3.6. Validating the predictive value of CRS

The 6 PGs were tested via qRT-PCR in the liver tissues of validation cohort (n=143). Values of CRS, FRS and of 6 genes based on mRNA expression were examined. Area under receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were used to evaluate diagnostic power of CRS.

3.7. Statistical analysis

Continuous variables and proportions were compared using Mann-Whitney and χ^2 Fisher's exact test, respectively. Univariate and multivariate analyses were performed using the Cox model. All statistical analysis was done using SPSS 19.0 (SPSS Inc, Chicago). $P<0.05$ was considered statistically significance.

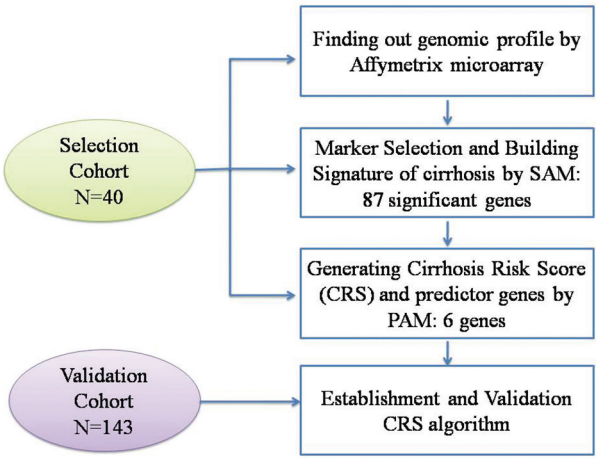

Figure 1. Work flow.

Table 1. Genes screened by SAM and PAM

Group	Significant genes	Predictor genes (threshold)
S3-4 versus S0	104	18 (4)
S3-4 versus S0-2	373	30 (5.5)
S1-4 versus S0	828	36 (5.5)
S0 versus S1-2 versus S3-4	610	46 (2)
In Summary	87	14

4. RESULTS

4.1. Characteristics of the patients

As the first step in this study, all genes significant in selection cohort (N=40, S0; S1; S2; S3; S4: n=8) from the genomic scan were identified. Next, the selection cohort was enrolled for signature building. To ensure the independent validation of the signature, the validation cohort (N=143, S0: n=39; S1: n=21; S2: n=34; S3: n=24; S4: n=25) was enrolled. Figure 1 illustrates the overall workflow and sample sources. The risk factors, such as age, sex, and fibrosis score, fibrosis rate, alanine aminotransferase (ALT), total bilirubin (TBil), albumin (Alb) did not differ significantly between the two cohorts. However, duration of infection, HBV DNA and the percentage patient with positive HBeAg were significantly higher in the Validation cohort.

4.2. Establishment of differentially expressed gene profile

To establish the gene expression profiles and their correlation with the histological fibrotic stages, a total of 1674 differentially expressed genes were identified in selection cohort (data not shown).

4.3. Selecting significant genes (SGs) and building genomic signatures of cirrhosis

SAM was able to detect differentially expressed genes in selection cohort. SGs based on several clinical endpoints were selected for building the signature. They were divided into 4 categories: (1) 104 SGs associated with bridging-fibrosis/cirrhosis when compared with no-fibrosis (S3-4 versus S0); (2) 373 SGs associated with bridging fibrosis/cirrhosis when compared with no portal/periportal fibrosis (S3-4 versus S0-2); (3) 828 SGs associated with any fibrosis when compared with no-fibrosis (S1-4 versus S0); (4) 610 SGs in fibrosis trend analysis (S0 versus S1-2 versus S3-4). The differential expression of all SGs found by SAM and the number of SGs is summarized in Table 1. To define the optimal set of genes able to delineate cirrhosis, we converged 678 genes identified by SAM tests. The expression pattern of these SGs had a good signature in analysis of unsupervised hierarchical clustering. Three main clusters were formed (S0, S1-S3 and S4) and the different cluster algorithms led to similar results which supported the gene expression profiles, robustly reflecting their histological classification. For the primary endpoints (category 1-2) and the other endpoints (category 3-4), 87 SGs duplicated in the above 2 kinds of comparison categories were included for further analysis.

4.4. Generating final CRS composed of 6 predictor genes (PGs)

For building CRS signature, PAM was done in selection cohort. All 87 duplicated SGs were enrolled to build the CRS signature and also analyzed in 4 categories by PAM. It was optimized by testing different thresholds; an approach that led to a reduced number of genes required for a robust prediction. The number of PGs found in PAM was also displayed (Table 1). It was seen that 90% of analyzed samples had a probability of almost 1 (meaning 100% fit), whereas the remaining samples had a suboptimal probability and thus did not fit the classification. Determination of whether a set of PGs had “Cirrhosis Risk” feature was done by including genes demonstrated to be highly significantly differential among the selection cohort. In summary, 14 PGs were found to be duplicated in the above 2 kinds of comparison categories to define the progression of cirrhosis. Then 14 PGs were quantified by use of qRT-PCR in liver tissues of selection cohort. Due to the fact that COL1A1/2 were already identified or reported in association with liver fibrosis, we chose the other 12 novel genes to elaborate the relationship between liver fibrotic stages. Six of them were, indeed, differentially expressed among fibrotic stages S4, S1-3 and S0 ($p \leq 0.0.5$) and shown to be significantly up-regulated in S4 compared to S0 in liver tissues (Figure 2). The genomic regulation was correlated best with cirrhosis. So we defined these 6 genes as having predicting “Cirrhosis Risk” signature. Thus a set of 6 novel genes

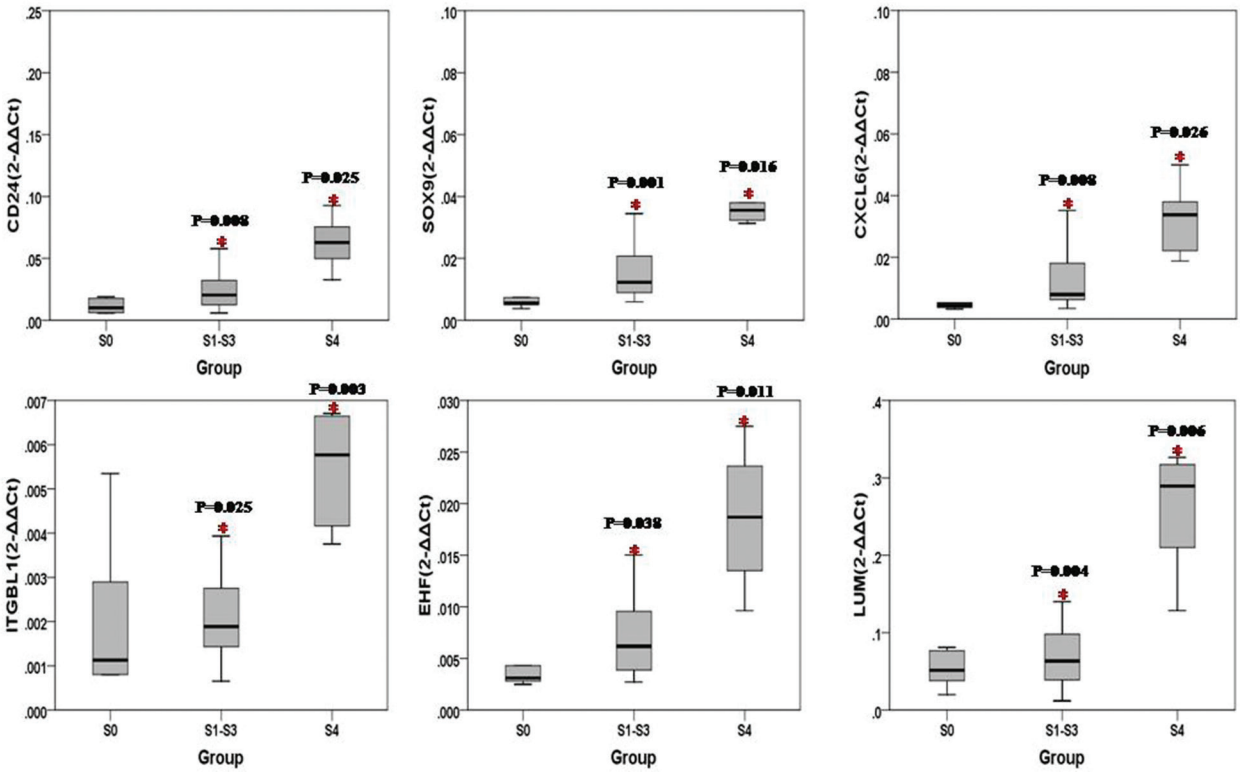

Figure 2. Six predictor genes identified by qRT-PCR. Six genes were elaborated in relationship with fibrotic stages in 40 patients of selection cohort. The box plots were shown for each gene. They were differentially expressed among S4, S1-3 and S0 (P≤0.05).

including CD24, CXCL6, EHF, ITGBL1, LUM and SOX9 could predict the progression of cirrhosis.

4.5. Establishment and validation CRS algorithm

The expression of 6 genes were further examined by qRT-PCR in liver tissues of validation cohort (n=143). The value of CRS and FRS based on mRNA expression levels of 6 genes with hepatic fibrotic stage was assessed utilizing logistic multivariable linear regression. AUROCs were established to evaluate diagnostic powers of CRS, FRS, 6 genes and clinic risk factors in liver fibrosis. We established 3 models including two CRS and one FRS. Values of AUROC, 95% CI, cut-off value, sensitivity and specificity are shown in Table 2. CRS was calculated based on 6 genes in each patient, hence reflecting the combined impact of all 6 genes. The cut-off value of CRS was 7.794, the higher the CRS value, the higher the risk is. The AUROC of CRS1 was 0.944 (95% CI: 0.87-1.00, P<0.001) for predicting the risk of developing cirrhosis. AUROC of CRS1 model distinguishing cirrhosis (fibrosis stages S4 from S0-S3) was higher than CRS2, 6 genes and clinic risk factors (Figure 3A/B). Sensitivity and specificity of CRS1 were significantly superior to others. AUROC of FRS distinguishing fibrosis (S0 from S1-S4) was also the highest among them (Figure 3C/D), but not ideal for distinguishing early stage of liver fibrosis. Taken together, the results clearly indicated that the CRS1 was a much better predictor of cirrhosis risk than clinical risk factors. These findings open a spectrum of genomic molecular predictors for cirrhosis.

4.6. Predictive value of CRS

Table 3 shows two cutoff values to identify CHB patients with low risk (<6.43) versus high risk (>8.32) of developing cirrhosis in the validation set. Additional cutoffs and their diagnostic values are calculated (unshown). A low cutoff value of <6.43 to identify low-risk patients would correctly classify 36 of 39 low-risk patients (specificity=58.6%). More importantly, the presence of high-risk patients could be excluded with great certainty because only 4 of 49 of the high-risk patients would fall into this category (misclassification rate=8.16%). Of the 40 patients with CRS1<6.43, 36 low-risk patients were correctly identified (NPV=98.5%). A high cutoff value of >8.32 to identify high-risk patients would correctly classify 13 of 49 high-risk patients (sensitivity=50%), and misclassify 0 of 39 low-risk patients (misclassification rate=0%). Of the 13 patients with CRS1>8.32, 13 high-risk patients were correctly identified (PPV=83.3%). Also, 36 (39.8%) of the patients fell between the 6.43 and 8.32, and hence could not be classified accurately (Table 3).

Table 2. Diagnostic values of CRS and others

CRS (S4 vs S0-3)	AUROC	95% CI		Cut-off value	Sen	Spe	FRS (S1-4 vs S0)	AUROC	95% CI		Cut-off value	Sen	Spe
CRS1	0.944	0.87	1.00	7.794	0.937	0.900	FRS	0.853	0.79	0.92	10.416	0.975	0.667
CRS2	0.849	0.76	0.94	0.190	0.631	1.000							
CD24	0.823	0.83	0.93	9.601	0.892	0.900	CD24	0.812	0.73	0.89	7.907	0.850	0.691
CXCL6	0.809	0.77	0.89	9.901	0.892	0.900	CXCL6	0.831	0.81	0.94	7.446	0.725	0.889
EHF	0.832	0.81	0.95	11.009	0.946	0.800	EHF	0.822	0.81	0.93	8.766	0.925	0.716
ITGBL1	0.805	0.70	0.89	7.881	0.838	1.000	ITGBL1	0.824	0.75	0.90	7.094	0.900	0.716
LUM	0.881	0.83	0.94	13.230	0.982	0.900	LUM	0.797	0.72	0.88	11.560	0.900	0.605
SOX9	0.870	0.75	0.99	8.952	0.811	0.900	SOX9	0.819	0.74	0.90	7.614	0.775	0.802
Age	0.798	0.69	0.90	39.500	0.532	1.000	Age	0.681	0.58	0.78	35.500	0.575	0.741
ALT	0.709	0.55	0.87	73.000	0.577	0.800	ALT	0.704	0.61	0.80	60.000	0.750	0.617
AST	0.766	0.64	0.89	84.500	0.829	0.700	AST	0.683	0.58	0.78	37.000	0.625	0.716
TBil	0.754	0.56	0.95	21.985	0.829	0.700	TBil	0.660	0.56	0.76	14.350	0.750	0.580
Abbreviations: Sen, sensitivity; Spe, specificity													

Table 3. Predictive values of CRS1

CRS1 values	Low risk S0 N=39	High risk S3-4 N=49	Sen (%)	Spe (%)	PPV (%)	NPV (%)	Misclassifying rate (%)	No. (%) patients	Interpretation
<6.43	36	4	90	58.56		98.50	8.16	40 (45.5)	Low risk for cirrhosis
6.43-8.32	3	32						35 (39.8)	Indeterminate
>8.32	0	13	50	99.10	83.30		0	13 (14.8)	High risk for cirrhosis
Abbreviations: Sen, sensitivity; Spe, specificity; PPV, positive predictive value; NPV, negative predictive value. Sensitivity and Specificity were calculated for a cutoff of<compared with >=									

So CRS1 could be considered to have good predictive value of progression to cirrhosis.

5. DISCUSSION

Despite the dramatic improvement in the management of CHB, the prognosis is extremely different between hepatitis and cirrhosis. Cirrhosis and HCC remain the major cause of morbidity and mortality, with the highest rates observed in Asia (12-13). Natural history studies in untreated patients have reported annual HCC incidences of 0.3-0.6% in non-cirrhotic patients, increasing to 2.2-3.7% in compensated cirrhotic patients (13). According to Europe, America and Asian-Pacific consensus statement (14-15), antiviral therapy is recommended for patients with an increased risk of developing cirrhosis. Entecavir is currently recommended as first-line options for the treatment of CHB, which maintains long-term viral suppression in over 95% of patients and reverses histologic cirrhosis (16-17). The lack of a prognostic test to identify patients at high-risk for cirrhosis resulted in the current recommendation of treatment: candidates as those with portal or bridging-fibrosis on a liver biopsy. While this approach will identify those with significant disease, it does not indicate the likelihood of developing cirrhosis in those patients with less severe histology. Therefore, early detection and treatment for patients at high-risk of cirrhosis are important.

The CRS reported here stratifies the cirrhosis risk in Chinese CHB patients. For the first time, one can estimate the risk of cirrhosis rather than using findings of the liver biopsy to project the future course of disease. Liver biopsy represents only a one time point in the long natural history of CHB, whereas genetic markers are intrinsic and “life-long”. Potentially, CRS could be used to stratify patients’ cirrhosis risk prior to liver biopsy. Consistent with the recommendations, one can perhaps make the argument that patients at high risk should be treated regardless of their fibrosis stage. Patients with low or indeterminate risk could then undergo liver biopsy and those with fibrosis stage S2-4 would also be treated. Among those with fibrosis stage S0-1, treatment could probably be deferred in those at low risk and individualized in those with indeterminate risk. Such a strategy would improve current management in two aspects. First, liver

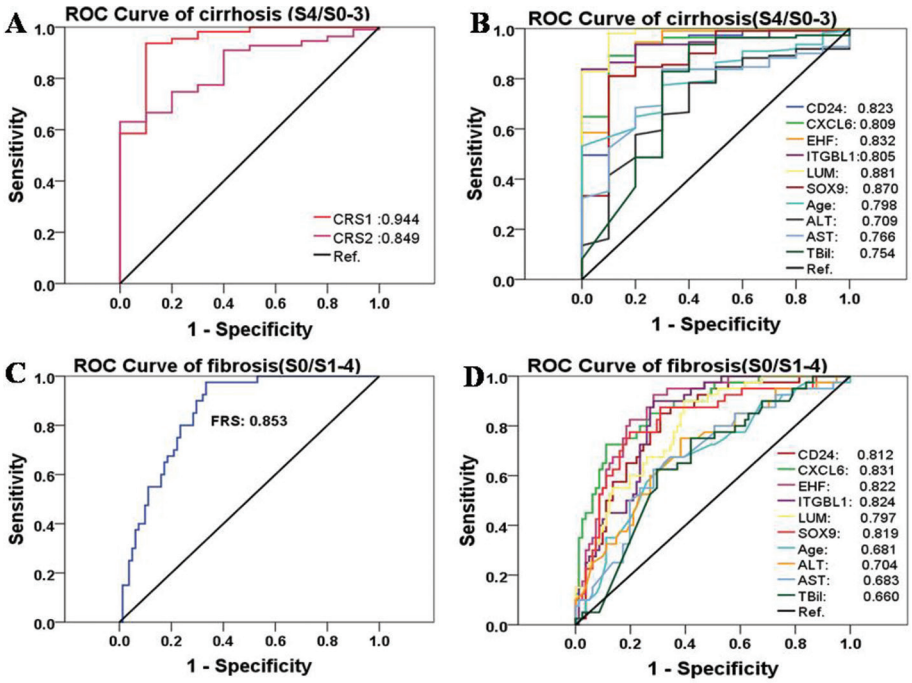

Figure 3. Diagnostic accuracy of prediction of liver fibrosis and cirrhosis. ROC curves for CRS1/2, FRS, 6 genes and clinical factor values were used to predict fibrosis and cirrhosis. The area under the ROC curve is depicted in parentheses. Diagnostic values of CRS1/2 (A), 6 genes and clinical factors (B) for detecting cirrhosis are shown. Diagnostic values of FRS (C), 6 genes and clinical factors (D) for detecting fibrosis are shown.

biopsy could perhaps be avoided in patients at high risk of developing cirrhosis. Second, those patients with fibrosis stage S0-1 but at high risk for cirrhosis, whose treatment is currently deferred, might benefit from immediate treatment as their response to antiviral therapy would be higher. On the other hand, those patients at high risk but who decided to defer antiviral therapy or failed a prior course of antiviral therapy may need closer monitoring and more aggressive management compared to low-risk patients.

Although several studies had previously described the gene expression pattern of a subgroup of lesions in other etiologies related to liver fibrosis (18-21). The goal of this study was to build a signature with a minimal set of highly predictive genes based on an exhaustive list of significant markers. Most of the PGs had confirmed associations with certain clinical endpoints, but only 6 formed the final CRS signature. It is important to emphasize that the entire signature building process, including selecting source genes, ranking genes, and building the final signature, was performed strictly within the selection cohort. The validation cohort was only used to validate the final signature once it was established. We expect the CRS results would be applicable to other independent sample sets, and the signature-building approach described here can be generalized to build predictive signatures in other diseases. In contrast, several clinical factors (age, ALT, AST and TBil) also had associations with cirrhosis risk. Moreover, the CRS measurements are objective, whereas clinical factors are subjective and suffer from recall bias and inaccuracy. Taken together, these reasons explained why CRS is a better predictor than the clinical factors studied here.

Complex diseases such as fibrosis associated with CHB are likely polygenic. Indeed, each of the 6 most predictive markers provided only moderate predictability, whereas the combination of these 6 was more robust and predictive. This observation raised the question of whether the presence of 6 novel genes (CD24, CXCL6, EHF, ITGBL1, LUM and SOX9) that correlated more extensively with cirrhosis, would point towards a higher incidence of cirrhosis or whether they have a detrimental impact on liver function. It was indicated that human CD24 gene polymorphisms may affect both the risk and the progression of chronic HBV infection (22). CXC chemokines can be subclassified into ELR^+ or ELR^- based on the presence of a tripeptide motif ELR (Glu-Leu-Arg) at the NH^2 terminus. CXCL6 is a member of the ELR^+ CXC family and high expression levels of CXCL6 are associated with a worse prognosis and increased severity of portal hypertension in patients with alcoholic hepatitis (23). The transcription factor, sex-determining region Y-box 9 (SOX9), becomes ectopically expressed in activated HSCs, where it is responsible for COL1 production. During development, SOX9 has diverse roles regulating the expression of a number of genes encoding ECM proteins. SOX9 has also been associated with fibrotic pathologies affecting the skin,

kidney, and heart (24-25). Pritchett demonstrated that SOX9, downstream of Hedgehog signaling, is a core factor mediating the expression of ECM components involved in liver fibrosis (26). Others including EHF, ITGBL1 and LUM have not been found to participate in liver fibrosis.

In conclusion, we have demonstrated that a CRS signature containing 6 predictive genes can identify patients with CHB at high risk for cirrhosis. Relying on cutoff values such as 6.43 and 8.32, we could distinguish between the absence and presence of high cirrhosis risk with sufficient reliability. Application of CRS in clinical practice could help to reduce the rate of liver biopsy in patients with CHB, identify additional candidates for treatment at early disease stage, and assist re-treatment decisions.

6. ACKNOWLEDGMENT

Ming-Yi Xu and Ying Qu contributed equally to this work. This study was supported by the Development Program of China during the 12th Five year Plan Period (No.2012ZX10002007-001-040 & 2013ZX10002004-002-003), the Science and Technology Commission of Shanghai Municipality (No. 12DZ1941603) and Outstanding Young Physician Training Program of Shanghai First People's Hospital (No.061405). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

7. REFERENCES

1. NIH consensus development statement on management of hepatitis B. *NIH Consens State Sci Statements* 25, 1-29 (2008)
2. RJ Fontana: Management of patients with decompensated HBV cirrhosis. *Semin Liver Dis* 23, 89-100 (2003)
DOI: 10.1055/s-2003-37591
3. P Marcellin, T Asselah, N Boye: Fibrosis and disease progression in hepatitis C. *Hepatol* 36, S47-S56 (2002)
DOI: 10.1002/hep.1840360707
4. MY Xu, Y Wang, Y Qu, YW Dong, LG Lu: Relationship of differential gene expression profiles to HBV-related fibrogenesis and carcinogenesis. *J Hepatol* 56, S415-S416 (2012)
5. MY Xu, Y Qu, YW Dong, Y Wang, XB Cai, LG Lu: Genomic profile and functional characterization of fibrosis in patients with chronic hepatitis B. *J Hepatol* 54, S135-S136 (2011)
6. PJ Scheuer: The nomenclature of chronic hepatitis: time for a change. *J Hepatol* 22, 112-114 (1995)
DOI: 10.1016/0168-8278(95)80269-X
7. PJ Scheuer: Classification of chronic viral hepatitis: a need for reassessment. *J Hepatol* 13, 372-374 (1991)
DOI: 10.1016/0168-8278(91)90084-O
8. H Huang, ML Shiffman, S Friedman S, R Venkatesh, N Bzowej, OT Abar, CM Rowland, JJ Catanese, DU Leong, JJ Sninsky, TJ Layden, TL Wright, T White, RC Cheung: A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. *Hepatol* 46, 297-306 (2007)
DOI: 10.1002/hep.21695
9. GW Wright, RM Simon: A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* 19, 2448-2455 (2003)
DOI: 10.1093/bioinformatics/btg345
10. L Bullinger, K Döhner, E Bair, S Fröhling, RF Schlenk, R Tibshirani, H Döhner, JR Pollack: Use of gene expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 350, 1605-1616 (2004)
DOI: 10.1056/NEJMoa031046
11. Tibshirani R, Hastie T, Narasimhan B, G Chu: Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 99, 6567-6572 (2002)
DOI: 10.1073/pnas.082099299
12. R Lozano, M Naghavi, K Foreman, S Lim, K Shibuya, V Aboyans, J Abraham, T Adair, R Aggarwal, SY Ahn, et al: Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2095–2128 (2012)
DOI: 10.1016/S0140-6736(12)61728-0
13. G Fattovich, F Bortolotti, F Donato: Natural history of chronic hepatitis B: special emphasis on disease progression and prognostic factors. *J Hepatol* 48, 335–352 (2008)
DOI: 10.1016/j.jhep.2007.11.011
14. European Association For The Study Of The Liver: EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J Hepatol* 57, 167–185 (2012)
DOI: 10.1016/j.jhep.2012.02.010

15. YF Liaw, N Leung, JH Kao, T Piratvisuth, E Gane, KH Han, R Guan, GK Lau, S Locarnini: Chronic Hepatitis B Guideline Working Party of the Asian-Pacific Association for the Study of the Liver: Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2012 update. *Hepatol Int* 6, 531–561 (2012)

16. TT Chang, CL Lai, YS Kew, SS Lee, HS Coelho, FJ Carrilho, F Poordad, W Halota, Y Horsmans, N Tsai, H Zhang, DJ Tenney, R Tamez, U Iloeje: Entecavir treatment for up to 5 years in patients with hepatitis B e antigen-positive chronic hepatitis B. *Hepatol* 51, 422–430 (2010)
DOI: 10.1002/hep.23327

17. TT Chang, YF Liaw, SS Wu, E Schiff, KH Han, CL Lai, R Safadi, SS Lee, W Halota, Z Goodman, YC Chi, H Zhang, R Hindes, U Iloeje, S Beebe S, B Kreter: Long-term entecavir therapy results in the reversal of fibrosis/cirrhosis and continued histological improvement in patients with chronic hepatitis B. *Hepatol* 52, 886–893 (2010)
DOI: 10.1002/hep.23785

18. P Sancho-Bru, J Altamirano, D Rodrigo-Torres, M Coll, C Millán, J José Lozano, R Miquel, V Arroyo, J Caballería, P Ginès, R Bataller: Liver progenitor cell markers correlate with liver damage and predict short-term mortality in patients with alcoholic hepatitis. *Hepatol* 55, 1931-1941 (2012)
DOI: 10.1002/hep.25614

19. SS Kdalid, S Hamid, AA Siddiqui, A Qureshi, N Qureshi: Gene profiling of early and advanced liver disease in chronic hepatitis C patients. *Hepatol Int* 5, 782-788 (2011)

20. W Ahmad, B Ijaz B, S Hassan: Gene expression profiling of HCV genotype 3a initial liver fibrosis and cirrhosis patients using microarray. *J Trans Med* 10, 41-58 (2012)
DOI: 10.1186/1479-5876-10-41

21. M Moreno, JF Chaves, P Sancho-Bru, F Ramalho, LN Ramalho, ML Mansego, C Ivorra, M Dominguez, L Conde, C Millán, M Marí, J Colmenero, JJ Lozano, P Jares, J Vidal, X Forns, V Arroyo, J Caballería, P Ginès, R Bataller: Ghrelin attenuates hepatocellular injury and liver fibrogenesis in rodents and influences fibrosis progression in humans. *Hepatol* 51, 974-985 (2010)
DOI: 10.1002/hep.23421

22. DL Li, LH Zheng, L Jin, Y Zhou, H Li, J Fu, M Shi, P Du, L Wang, H Wu, GY Chen, P Zheng, Y Liu, FS Wang, S Wang: CD24 polymorphisms affect risk and progression of chronic hepatitis B virus infection. *Hepatol* 50, 735-742 (2009)
DOI: 10.1002/hep.23047

23. M Dominguez, R Miquel, J Colmenero, M Moreno, JC García-Pagán, J Bosch, V Arroyo, P Ginès, J Caballería, R Bataller: Hepatic expression of CXC chemokines predicts portal hypertension and survival in patients with alcoholic hepatitis. *Gastroenterology* 136, 1639-1650 (2009)
DOI: 10.1053/j.gastro.2009.01.056

24. K Piper Hanley, F Oakley, S Sugden, DI Wilson, DA Mann, NA Hanley: Ectopic SOX9 mediates extracellular matrix deposition characteristic of organ fibrosis. *J Biol Chem* 283, 14063-14071 (2008)
DOI: 10.1074/jbc.M707390200

25. J Pritchett, V Athwal, N Roberts, NA Hanley, KP Hanley: Understanding the role of SOX9 in acquired diseases: lessons from development. *Trends Mol Med* 17, 166-174 (2011)
DOI: 10.1016/j.molmed.2010.12.001

26. J Pritchett, E Harvey, V Athwal, A Berry, C Rowe, F Oakley, A Moles, DA Mann, N Bobola, AD Sharrocks, BJ Thomson, AM Zaitoun, WL Irving, IN Guha, NA Hanley, KP Hanley: Osteopontin is a novel downstream target of SOX9 with diagnostic implications for progression of liver fibrosis in humans. *Hepatol* 56, 1108-1116 (2012)
DOI: 10.1002/hep.25758

Key Words: Chronic Hepatitis B, CHB, Cirrhosis, Microarray

Send correspondence to: Lun-Gen Lu, Department of Gastroenterology, Shanghai First People's Hospital, Shanghai Jiao Tong University, No 100, Haining Road, Shanghai, 200080, China, Tel: 8621-63240090, Fax: 8621-66283869, E-mail: lungenlu1965@163.com