SOMAscan Proteomics Identifies Serum Biomarkers Associated With Liver Fibrosis in Patients With NASH

Yi Luo,[1] Samir Wadhawan,[1*] Alex Greenfield,[1*] Benjamin E. Decato,[1] Abdul M. Oseini,[2] Rebecca Collen,[2] Diane E. Shevell,[1] John Thompson,[1] Gabor Jarai,[1] Edgar D. Charles,[1] and Arun J. Sanyal[2]

Nonalcoholic steatohepatitis (NASH) is a major cause of liver-related morbidity and mortality worldwide. Liver fibrosis stage, a key component of NASH, has been linked to the risk of mortality and liver-related clinical outcomes. Currently there are no validated noninvasive diagnostics that can differentiate between fibrosis stages in patients with NASH; many existing tests do not reflect underlying disease pathophysiology. Noninvasive biomarkers are needed to identify patients at high-risk of NASH with advanced fibrosis. This was a retrospective study of patients with histologically proven NASH with fibrosis stages 0-4. The SOMAscan proteomics platform was used to quantify 1,305 serum proteins in a discovery cohort (n = 113). In patients with advanced (stages 3-4) versus early fibrosis (stages 0-2), 97 proteins with diverse biological functions were differentially expressed. Next, fibrosis-stage classification models were explored using a machine learning–based approach to prioritize the biomarkers for further evaluation. A four-protein model differentiated patients with stage 0-1 versus stage 2-4 fibrosis (area under the receiver operating characteristic curve [AUROC] = 0.74), while a 12-protein classifier differentiated advanced versus early fibrosis (AUROC = 0.83). Subsequently, the model's performance was validated in two independent cohorts (n = 71 and n = 32) with similar results (AUROC = 0.74-0.78). Our advanced fibrosis model performed similarly to or better than Fibrosis-4 index, aspartate aminotransferase–to-platelet ratio index, and nonalcoholic fatty liver disease (NAFLD) fibrosis score–based models for all three cohorts. *Conclusion:* A SOMAscan proteomics-based exploratory classifier for advanced fibrosis, consisting of biomarkers that reflect the complexity of NASH pathophysiology, demonstrated similar performance in independent validation cohorts and performed similarly or better than Fibrosis-4 index, aspartate aminotransferase–to-platelet ratio index, and NAFLD fibrosis score. Further studies are warranted to evaluate the clinical utility of these biomarker panels in patients with NAFLD. (*Hepatology Communications* 2021;5:760-773).

Nonalcoholic steatohepatitis (NASH) is defined as the presence of more than 5% hepatic steatosis and inflammation with hepatocyte injury (e.g., ballooning) with or without fibrosis.[1] NASH has emerged as a major cause of liver-related morbidity and mortality worldwide.[2,3] The greater rate of mortality in affected individuals is related to cardiovascular events, liver-related outcomes, and cancers.[4] The liver-related mortality risk is driven largely by the development of cirrhosis, which is preceded by hepatic fibrosis, a hallmark of disease progression.[5-7] Identification of fibrosis stage

Abbreviations: ALT, alanine aminotransferase; APRI, AST-to-platelet ratio index; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; CRN, Clinical Research Network; FIB-4, fibrosis-4 index; FDR, false discovery rate; GDF-15, growth/differentiation factor 15; IGF-1, insulin-like growth factor 1; IGFBP, insulin-like growth factor-binding protein; IL18BP, interleukin-18-binding protein; LTBP4, latent transforming growth factor beta binding protein 4; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD activity score; NFS, NAFLD fibrosis score; NASH, nonalcoholic steatohepatitis; RFU, relative fluorescent unit; ROC, receiver operating characteristic; SAP, serum amyloid P; SHBG, sex hormone–binding globulin; T2DM, type-2 diabetes mellitus; VCAM1, vascular cell adhesion molecule 1; VCU, Virginia Commonwealth University.

Received August 5, 2020; accepted December 17, 2020.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1670/suppinfo.

**These authors contributed equally to this work.*

Supported by Bristol Myers Squibb.

© 2021 The Authors. Hepatology Communications published by Wiley Periodicals LLC on behalf of the American Association for the Study of Liver Diseases. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

therefore allows for the assessment of mortality risk and selection of individuals who need more intensive care, to prevent the progression to cirrhosis. Those with advanced fibrosis (i.e., NASH Clinical Research Network [CRN] stages 3 and 4) are at greatest risk of adverse outcomes, and are therefore a population of interest, as they are most likely to benefit from effective interventions.

The identification of advanced fibrosis in the overall population of patients with NASH is challenging. First, most patients are asymptomatic; when symptoms do occur, they are often nonspecific,[8] and when the signs of decompensated or advanced disease appear, therapeutic options are limited. Given the asymptomatic nature of the disease for long periods, diagnosis can be further delayed, as primary care providers are often the only point of contact for most affected individuals with relevant risk factors (e.g., type 2 diabetes mellitus [T2DM], obesity, and hypertension). The ability to identify individuals with NASH and, more specifically, those at greatest risk of adverse health outcomes remains a critical unmet need. An ideal solution should involve a test, such as a blood test, that can easily be performed in a primary care setting.

There are currently no validated noninvasive tests capable of distinguishing between different hepatic fibrosis stages in patients with nonalcoholic fatty liver disease (NAFLD) or NASH. There are, however, a large number of simple diagnostic aids that do not require special tests, such as the aspartate aminotransferase (AST)–to-platelet ratio index (APRI) and the fibrosis-4 (FIB-4) tests, as well as a number of special biomarker panels that are being developed. The latter include the Enhanced Liver Fibrosis test (Siemens Healthineers, Erlangen, Germany), FibroSure (in the United States; known as FibroTest outside the United States; LabCorp, Burlington, NC), FibroMeter (ARUB Laboratories, Salt Lake City, UT), and a panel based on collagen fragments released during fibrogenesis; these have been reviewed elsewhere.[9] Unfortunately, none of these tests are currently approved by regulatory agencies; all have limitations and do not reflect the complexity of NASH pathophysiology. There is also considerable progress in magnetic resonance imaging–based methods to assess fibrosis, which require additional visits to imaging centers.[10] Similarly, while transient elastography-based and ultrasound-based methods are also being developed, the specificity for diagnosis of discrete fibrosis stages is modest.[10] Thus, there is a continued unmet need to define biomarkers or biomarker panels that can be used to identify fibrosis stage in those patients with clinical risk factors for NAFLD.

Proteomic analysis allows the simultaneous assessment of many proteins within a biological sample. The technological platform for targeted proteomic analysis of serum is evolving and several groups have developed tools that allow scaling for broader deployment. SOMAscan (SomaLogic, Inc., Boulder, CO) involves a unique protein measurement system using Slow Off-Rate Modified Aptamer (SOMAmer) molecules that bind to proteins with high affinity and

View this article online at wileyonlinelibrary.com.
DOI 10.1002/hep4.1670

Potential conflict of interest: Dr. Luo owns stock in and is employed by BMS. Dr. Wadhawan owns stock in BMS. Dr. Decato is employed by BMS. Dr. Shevell owns stock in and is employed by BMS. Dr. Jarai owns stock in BMS. Dr. Charles owns stock in BMS and Merck. He is employed by BMS. Dr. Sanyal consults and received grants from Conatus, Gilead, Echosens-Sandhill, Mallinckrodt, Salix, Novartis, Galectin, and Sequana. He consults and owns stock in GenFit, Hemoshear, Durect, and Indalo. He consults for Immuron, Intercept, Pfizer, Boehringer Ingelheim, Nimbus, Lilly, Merck, Novo Nordisk, Fractyl, Allergan, Chemomab, Affimmune, Teva, Ardelyx, Terns, ENYO, Birdrock, Albireo, Sanofi, Takeda, Janssen, Zydus, BASF, AMRA, Perspectum, Owl, Poxel, Servier, Second Genome, General Electric, and 89Bio. He received grants from BMS. He received royalties from Elsevier and Uptodate. He owns stock in Exhalenz, Akarna, and Tiziana. He is employed by Sanyal Bio.

ARTICLE INFORMATION:

From the [1]Bristol Myers Squibb, Princeton, NJ, USA; [2]Virginia Commonwealth University, Richmond, VA, USA.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Yi Luo, Ph.D.
Bristol Myers Squibb
3401 Princeton Pike
Lawrenceville, NJ 08648
E-mail: yi.luo@bms.com
Tel.: +1-609-302-5675

specificity.[11] SOMAscan has been used to identify diagnostic signatures of several other diseases.[12-14] In this study, the SOMAscan proteomics approach was used to identify biomarkers that distinguish fibrosis stages, particularly advanced fibrosis, in a population of patients with NASH.

Methods

This was a retrospective study performed on stored serum samples from a cohort of patients with histologically proven NASH of varying severity and fibrosis stage according to NASH CRN criteria.[15] All patients were recruited at a single center. These patients were part of a longitudinal natural history study. The study was approved by the institutional review board (Virginia Commonwealth University [VCU], 1960), and all patients provided informed consent at the time of enrollment. The samples were analyzed with funding support from the sponsor (Bristol Myers Squibb, Princeton, NJ), and the raw data processing and analysis were performed jointly by investigators from Bristol Myers Squibb and VCU. The authors have written the manuscript and take responsibility for the contents of this manuscript. Transparent reporting of individual prognosis or diagnosis (TRIPOD) standards for reporting of biomarker studies were met (Supporting Table S1).

STUDY DESIGN AND CONTEXT OF USE

The context of use was to develop and validate a proteomics-based diagnostic signature of NASH with advanced fibrosis (NASH CRN stages 3-4) in patients with histologically proven steatohepatitis. This was done by an initial unbiased analysis of stored serum samples from a histologically phenotyped cohort of individuals with NASH. The diagnostic signature was next validated in two independent validation cohorts.

PATIENT POPULATION

Discovery Cohort

This cohort was used initially for discovery of a diagnostic signature for fibrosis stages in patients with NASH. Originally, these samples were obtained as part of a study on the natural history of NAFLD with the aim of evaluating circulating factors associated with changes in disease phenotype over time. Individuals with suspected NAFLD who were being considered for a liver biopsy as part of standard of care were screened for this study. Individuals who provided informed consent were included in this study if they had histologically proven NASH according to NASH CRN criteria.[15] The "nonalcoholic" nature of the disease was established by the Alcohol Use Disorders Identification Test questionnaire, which included consumption of less than 2 units of alcohol daily for women and 3 units of alcohol daily for men.[16] All available samples from the natural history study were analyzed by SOMAscan proteomics and have not been used in other proteomics studies.

Validation Cohort

The diagnostic signature for advanced fibrosis was validated in two independent cohorts. The first cohort consisted of individuals who had participated in a phase 2 multicenter clinical trial of pegbelfermin, a PEGylated fibroblast growth factor 21 analog (MB130-045; NCT02413372) for the treatment of NASH.[17] Eligible patients had liver biopsies performed within 1 year of the study that demonstrated steatohepatitis with NASH CRN stage 1-3 fibrosis based on local pathology interpretation; other causes of chronic liver disease were ruled out in the sample donors.[17] The second cohort was also from a natural history study from the same institution where the discovery cohort was derived (VCU), and the patients met the same histologic diagnostic criteria as the discovery cohort.

HISTOLOGICAL ASSESSMENT OF LIVER DISEASE IN SAMPLES FROM VCU-DERIVED COHORTS

The reference standard for the assessment of hepatic fibrosis was histological stage of the liver. This was ascertained by a percutaneous or transjugular liver biopsy followed by histological examination of paraffin-embedded liver tissue sections. Sections were stained with hematoxylin and eosin as well as with Masson's trichrome stain. All histological samples were assessed by a single pathologist, after which another pathologist adjudicated each finding.

TABLE 1. PATIENT CHARACTERISTICS OF DISCOVERY COHORT

Parameters	Stage 0 (n = 13)	Stage 1 (n = 19)	Stage 2 (n = 36)	Stage 3 (n = 23)	Stage 4* (n = 22)
Male, n (%)	5 (38)	8 (42)	11 (30)	6 (26)	3 (14)
Age, median (Q1, Q3), years	44 (38, 57)	52 (45, 58)	56 (51, 61)	60 (50 , 62)	58 (51, 65)
BMI, median (Q1, Q3; n), kg/m^2	36.8 (32.2, 41.4; 2)	34.1 (33.8, 36.5; 7)	32.6 (28.9, 36.2; 8)	31.5 (28.4, 31.9; 5)	31.9 (29.2, 33.9; 5)
T2DM, n/N (%)	2/11 (18)	6/15 (40)	15/34 (44)	10/18 (56)	11/17 (65)
NAFLD activity score, median (Q1, Q3)	4 (3, 5)	4 (4, 4)	3.5 (3, 5)	4 (3.5, 5)	4 (3, 4.8)
ALT, median (Q1, Q3; n), U/L	65 (36, 90; 12)	60 (45, 77; 16)	47 (39, 79; 33)	74 (45, 106; 18)	46 (39, 52; 19)
AST, median (Q1, Q3; n), U/L	39 (30, 45; 12)	47 (37, 56; 16)	40 (29, 53; 33)	60 (41, 76; 18)	42 (39, 58; 19)
Platelet count, median (Q1, Q3; n), $\times 10^9$/L	230 (241, 279; 7)	220 (248, 290; 16)	272 (205, 321; 22)	199 (173, 278; 16)	177 (152, 210; 9)

Note: Fibrosis stage was assigned according to biopsy results. Number of patients (n) is shown for parameters for which patient data are missing.
*No patients with stage 4 fibrosis had decompensated cirrhosis.
Abbreviations: BMI, body mass index; Q, quartile.

Disease activity was scored using the NAFLD activity score (NAS).[15] Fibrosis staging was done using the NASH CRN classification system.[15]

SERUM ACQUISITION AND STORAGE

Blood samples were obtained on the day of the liver biopsy for patients in the VCU natural history studies; samples from study MB130-045 were obtained within 1 year of liver biopsy. Serum was separated and stored at −80°C within 1 hour of blood draw. Samples were maintained in storage until they were aliquoted for shipment to the reference lab. The samples underwent up to two freeze-thaws before being loaded onto the SOMAscan platform.

SAMPLE ANALYSIS

The SOMAscan assay was used to quantify the expression of 1,305 proteins in each serum sample. SOMAscan data underwent quality control and transformation based on bioinformatics standards.[18] Specifically, data were checked for batch effects with respect to experimental and demographic parameters using principle component analysis. Data were log2-transformed within each sample. Differential protein expression analyses were performed using the R-based LIMMA software package.[19] The data were adjusted for age and gender, as well as plate differences during sample analysis. The results were used to identify proteins significantly associated with fibrosis stage using a Benjamini-Hochberg false discovery rate (FDR) threshold of 1% and an absolute percent change of ≥25% between patients with NASH and advanced fibrosis (stage 3-4) versus early fibrosis (stage 0-2). A customized Luminex multiplex immunoassay was used to confirm differential expression of certain proteins indicated by SOMAscan. Luminex reagents were purchased from R&D Systems, Inc. (Minneapolis, MN), and assays were performed following manufacturer's instructions (see the Supporting Information Methods for additional assay details).

PREDICTIVE MODELING AND FEATURE SELECTION

Predictive models were built to address whether subsets of serum proteins could classify different stages of fibrosis. All 1,305 proteins measured by the SOMAscan array were used as input for the machine learning approach. The outcome of the model was fibrosis stage. The Elastic-Net algorithm was used with the multinomial link function (the rationale for use of Elastic-Net is described in the Supporting Information). The Elastic-Net algorithm was trained on the discovery cohort. To both fit the regularization parameter and calculate less-biased estimates for model performance, repeated runs of five-fold cross-validation were done. In addition, the discovery cohort model was validated using two independent test data sets from NASH cohorts: validation

cohort 1, patient baseline data from the MB130-045 trial; and validation cohort 2, patient data from a natural history study cohort. The model was considered validated if the different cohorts had areas under the receiver operating characteristic (AUROCs) similar to the discovery cohort.

Results

PATIENTS

A total of 113 individuals with NASH were included in the discovery cohort (Table 1). Most patients in the discovery cohort were female (70.8%), with a median age of 56.3 (range, 24.9 to 78.5) years, and 46.3% had T2DM. There were no meaningful differences in alanine aminotransferase (ALT), AST, body mass index, and NAFLD activity scores across fibrosis stages; however, patients with advanced fibrosis (stage 3-4) tended to be older and had lower platelet counts compared to patients with stage 0-2 fibrosis. Most patients from validation cohort 1 were female (63.4%), with a median age (range) of 51 (22-72) years, and 64.8% had T2DM (Supporting Table S2). Most patients in validation cohort 2 were female (65.6%), with a median age (range) of 50 (22-73) years, and 28.1% had T2DM (Supporting Table S3). The validation cohorts lacked patients with fibrosis stage 4.

IDENTIFICATION OF SERUM PROTEINS ASSOCIATED WITH FIBROSIS STAGES IN PATIENTS WITH NASH

Of the 1,305 proteins in the SOMAscan panel, 97 proteins were significantly differentially expressed in advanced (stage 3-4) versus early (stage 0-2) fibrosis, with an FDR threshold of 0.01 and an absolute percent difference of ≥25% (Fig. 1 and Supporting Table S4). Of these proteins, 90 had higher median levels in stage 3-4 patients compared with stage 0-2 patients; the remaining seven proteins had lower median levels in stage 3-4 patients compared with stage 0-2 patients. The identified proteins have known functions related to cell adhesion, extracellular matrix remodeling, innate and adaptive immune response, angiogenesis, cell stress response, and fibrogenesis. When comparing fibrosis stage 4 with stage 3, no protein met the FDR cutoff of 0.01; however, some proteins exhibited a trend of difference with raw *P* values < 0.01 (Supporting Table S5). No differentially expressed proteins were identified by comparing stage 1 versus stage 0 or stage 2 versus stage 1 or stage 0.

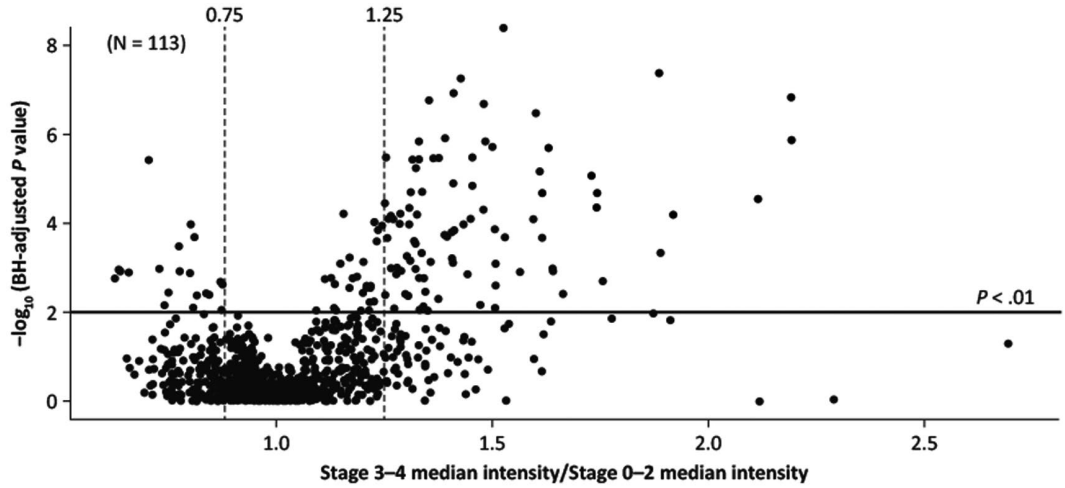

FIG. 1. Volcano plots for differentially expressed proteins in patients with stage 3-4 versus stage 0-2 fibrosis. Differentially expressed protein biomarkers in advanced (stage 3-4) versus early (stage 0-2) fibrosis were identified with FDR threshold of 0.01 and an absolute percentage difference of ≥25%. Abbreviation: BH, Benjamini-Hochberg.

BIOMARKER CLUSTERING ANALYSIS SUGGESTS MOLECULAR COMPLEXITY AND HETEROGENEITY OF PATIENTS WITH NASH AND ADVANCED FIBROSIS

A clustering analysis was performed on the 97 identified biomarkers to explore patient segments at the molecular level. Patients from this cohort were clustered into five groups (C1 through C5; Supporting Fig. S1). Most of the stage 4 patients were clustered into C1-C2, and stage 3 patients were clustered into C2-C3; most of the stage 0 and 1 patients were clustered in C4-C5. Stage 2 patients were the most heterogeneous group, clustered with both stage 3 and stage 0, and 1 patient in clusters C3-C5.

Further clustering analysis of only stage 3 and 4 patients revealed distinct subtypes referred to as S1 and S2 (Fig. 2). These subtypes displayed significant association with fibrosis stage ($P < 0.001$, Fisher's exact test), with 21 of 23 stage 3 patients belonging to subtype S1 and 13 of 22 stage 4 patients belonging to subtype S2. Of the 97 proteins used to cluster the samples, 62 showed significant differences between subtypes (FDR threshold of 0.01, absolute percent difference ≥25%; Fig. 3 and Supporting Table S6). Top proteins with higher abundance in S2 compared with S1 (S2/S1 median intensity > 2) included galectin-3-binding protein (LGALS3BP), insulin-like growth factor-binding protein 2 (IGFBP2), neurexin 3 (NRXN3), interleukin-18-binding protein (IL18BP), neural cell adhesion molecule (NCAM), and CXCL13, whereas proteins higher in abundance in S1 (S2/S1 median intensity < 0.75) included angiopoietin 1 (ANGPT1), insulin-like growth factor-binding protein 3 (IGFBP3), and insulin-like growth factor 1 (IGF-1).

We also examined the characteristics of S1 and S2 subtypes by exploring their associations with clinical parameters. Consistent with the enrichment of F4 in the S2 subtype, we observed significantly higher international normalized ratio and total bilirubin in

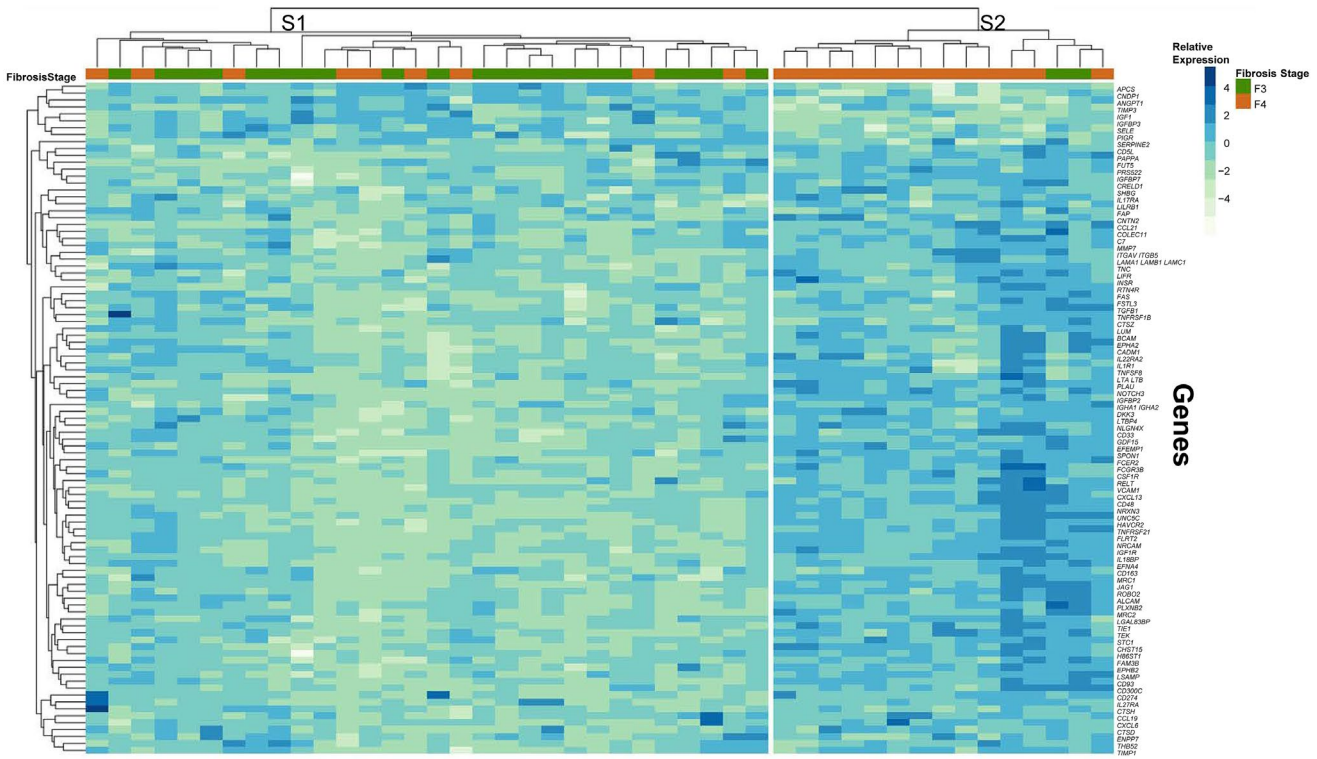

FIG. 2. Clustering analysis of 97 identified biomarkers in patients with stage 3 and 4 fibrosis. Heat map shows the unsupervised clustering of differentially expressed proteins in patients with NASH and advanced fibrosis (stages 3-4). Patients were grouped into two major subtypes: S1 and S2.

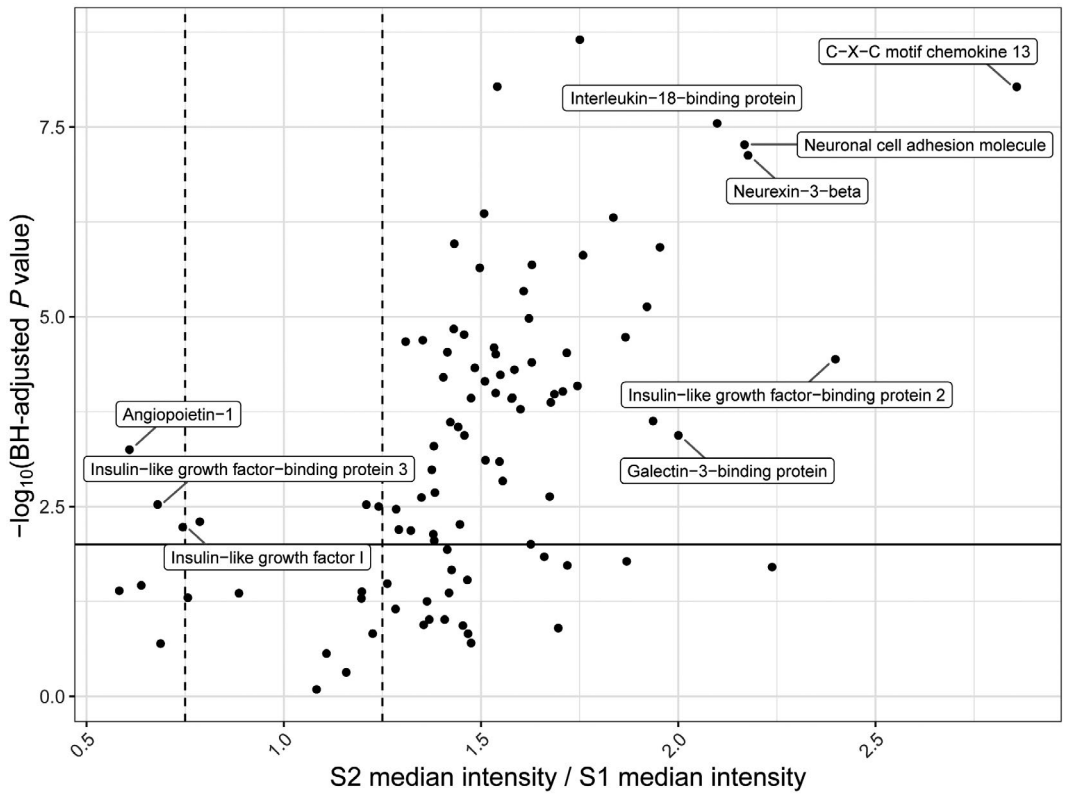

FIG. 3. Volcano plot for differentially expressed proteins in clusters S1 versus S2. Differentially expressed protein biomarkers in S1 versus S2 were identified with FDR threshold of 0.01 and an absolute percentage difference of ≥25% (dashed lines). Abbreviation: BH, Benjamini-Hochberg.

S2 compared with S1, and lower albumin and platelet counts in S2 compared with S1 (box plots and Benjamini-Hochberg-adjusted Wilcoxon rank-sum test *P* values provided in Supporting Fig. S2). No differences were observed in ALT, AST, or alkaline phosphatase, or associations of gender ($P = 1.0$, Fisher's exact test) or diabetes ($P < 0.11$, Fisher's exact test) with cluster status.

HORMONAL PATHWAYS ARE DYSREGULATED IN PATIENTS WITH NASH AND ADVANCED FIBROSIS

SOMAscan analysis levels of IGF-1 and its main carrier protein IGFBP-3 were lower, whereas IGFBP-1, IGFBP-2, IGFBP-7, and soluble insulin-like growth factor 1 receptor levels were elevated in patients with advanced fibrosis, resulting in a marked decrease in the ratios of IGF-1/IGFBP-1, IGF-1/IGFBP-2, and IGF-1/IGFBP-7 (Fig. 4 and Supporting Fig. S3). Moreover, sex hormone–binding globulin (SHBG), which regulates free sex hormone levels, was elevated in patients with advanced fibrosis compared to those patients with fibrosis stages 0-2, with highest median values observed in stage 4 (Supporting Fig. S3). SHBG levels were negatively correlated with IGF-1 (Supporting Fig. S4; Spearman $\rho = -0.42$ [$P < 0.0001$] and IGF-1/IGFBP-1 ratio [$\rho = -0.58$; $P < 0.0001$]).

CONFIRMATION OF DIFFERENTIALLY EXPRESSED PROTEINS

To confirm the identified biomarkers, available Luminex multiplex immunoassays were used to quantify the following biomarker hits from the same samples: growth/differentiation factor 15 (GDF-15), matrix metallopeptidase 7 (MMP-7), chitinase-3-like protein 1 (YKL-40), and tissue inhibitor of matrix

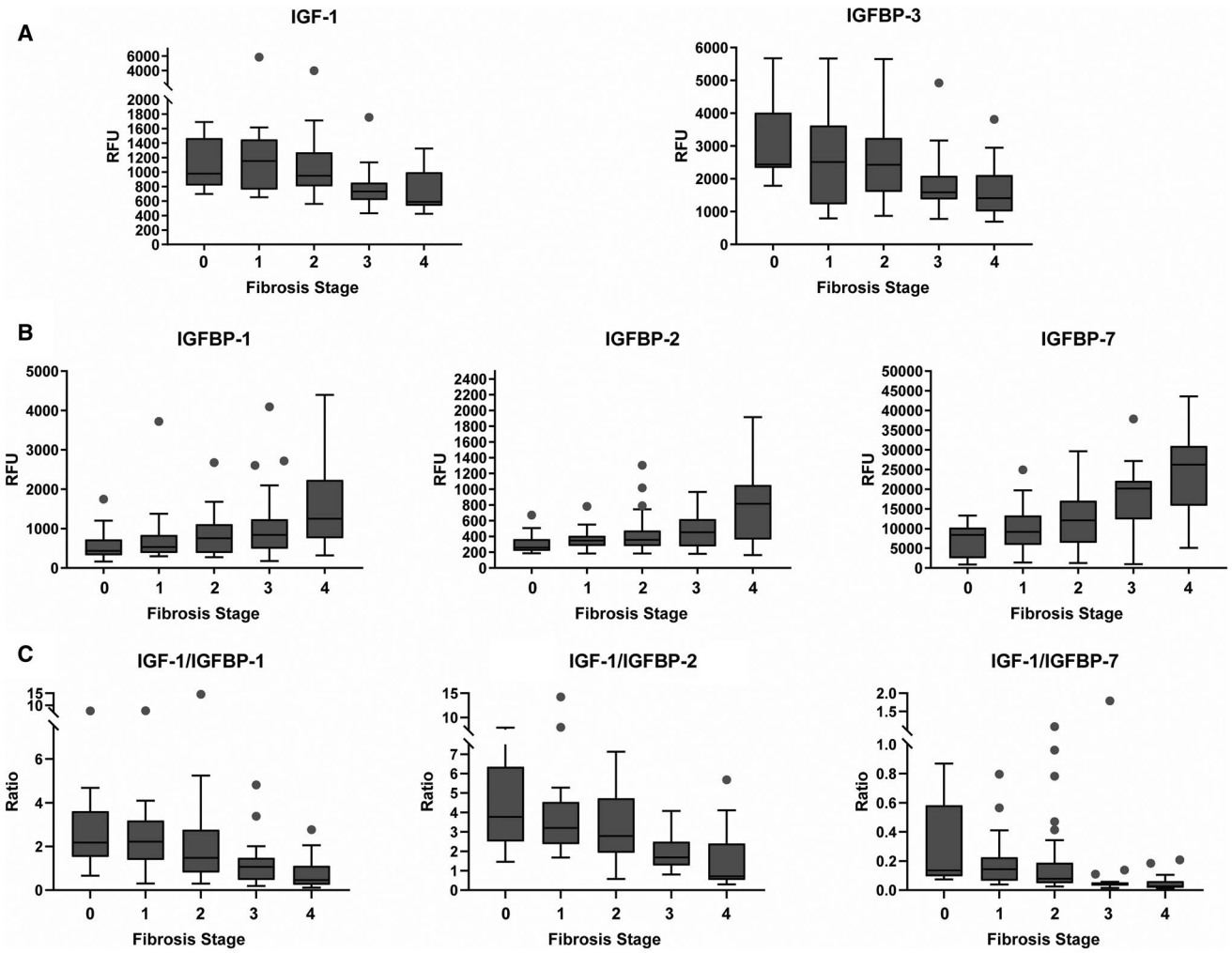

FIG. 4. Differential expression of IGF pathway proteins in patients with advanced fibrosis. SOMAscan data in RFUs are shown by fibrosis stage for IGF-1, IGFBP-3 (A); IGFBP-1, IGFBP-2, IGFBP-7 (B); and ratio for IGF-1/IGFBP-1 and IGF-1/IGFBP-7 (C). Data are depicted as box plots, in which the horizontal line shows the median value, box represents the interquartile range, and whisker and dots represent minimum and maximum values. Abbreviation: RFU, relative fluorescent unit.

metalloproteinase-1 (TIMP-1). The results from the Luminex assay were consistent with those obtained from SOMAscan and confirmed that these biomarkers are elevated in patients with NASH and advanced fibrosis compared to those with early fibrosis (Fig. 5).

MACHINE LEARNING APPROACH TO IDENTIFY BIOMARKER CLASSIFIERS FOR FIBROSIS STAGES

To select and prioritize the biomarkers identified in SOMAscan analysis, all 1,305 proteins quantified by SOMAscan were included in machine learning to identify models to classify patients with different fibrosis stages. The performance of the machine learning model in discriminating fibrosis stages is shown in Fig. 6. Four proteins (serum amyloid P [SAP], fibrinogen, olfactomedin, and SHBG) were identified in classifying stage 0-1 from stage 2-4 (mean AUROC [SD], 0.74 [0.026]). A total of 12 proteins (latent transforming growth factor beta binding protein 4 [LTBP4], IGF-1, vascular cell adhesion molecule 1 [VCAM1], interleukin-1 soluble receptor type-1, IL18BP, thrombospondin-2, collectin kidney 1, SHBG, interleukin-27 receptor subunit alpha, leukemia inhibitory factor receptor, soluble, fibulin-3, and plexin-B2) were selected in

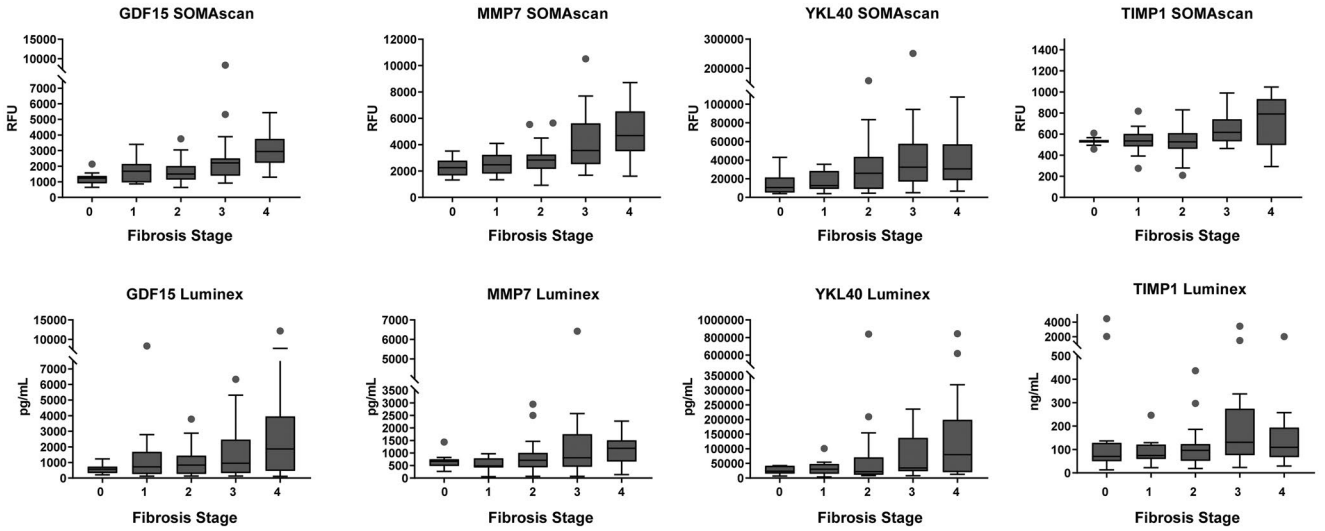

FIG. 5. Confirmation of differentially expressed serum proteins identified by SOMAscan using Luminex multiplex immunoassays. SOMAscan data in RFUs and immunoassay data in concentrations are shown for GDF-15, MMP-7, YKL-40, and TIMP-1. Data are depicted as box plots, in which the horizontal line shows the median value, box represents the interquartile range, and whisker and dots represent minimum and maximum values. Abbreviations: MMP7, matrix metallopeptidase 7; TIMP-1, tissue inhibitor of matrix metalloproteinase-1; YKL-40, chitinase-3-like protein 1.

classifying stage 3-4 from stage 0-2 (mean AUROC [SD], 0.83 [0.017]). No protein classifier was identified to discriminate stage 2 from stage 0-1 and stage 3 versus 4 (mean AUROC [SD], 0.61 [0.030]), which was consistent with the clustering analysis that showed patients with histologically assessed stage 2 fibrosis are very heterogeneous at the molecular level.

To validate the models, two independent NASH patient cohorts were identified and serum samples from these cohorts were tested using SOMAscan. The patient cohort from the pegbelfermin clinical trial MB130-045[17] was used as validation cohort 1, and validation cohort 2 was from a small natural history study. All patients from validation cohorts 1 and 2 had a NASH diagnosis using biopsy-based histology analysis. In both cohorts, ALT and AST values had no clear association with fibrosis stage. The stage 3-4 classifier demonstrated similar performance in both validation cohorts suggested by the similar AUROC (Fig. 7). Stage 0-1 classifier showed comparable performance only in validation cohort 1. To understand the performance of the SOMAscan-derived model in the context of existing indices used to identify fibrosis stages 3-4 such as FIB-4, APRI, and the NAFLD Fibrosis Score (NFS), these indices were calculated for the three study cohorts and AUROCs were compared. To classify fibrosis stages 3-4, the SOMAscan-derived model had AUROCs that were equal to or higher than those AUROCs generated using FIB-4, APRI, and NFS (Supporting Fig. S4).

Discussion

There remains a major unmet need to develop tests that can be used in a routine clinical setting to identify individuals with NASH who have advanced fibrosis. In this study, we used serum samples from characterized populations of patients with NASH and varying stages of fibrosis to identify biomarkers for patients with advanced fibrosis. This study provided data that indicated levels of a limited set of circulating proteins involved in diverse biological mechanisms may provide a signature for advanced fibrosis with relatively high fidelity.

The potential utility of a biomarker panel depends on analytical robustness, the biological plausibility of it being linked to the process being measured, sensitivity, specificity, and diagnostic performance across multiple cohorts of patients with a full spectrum of disease.

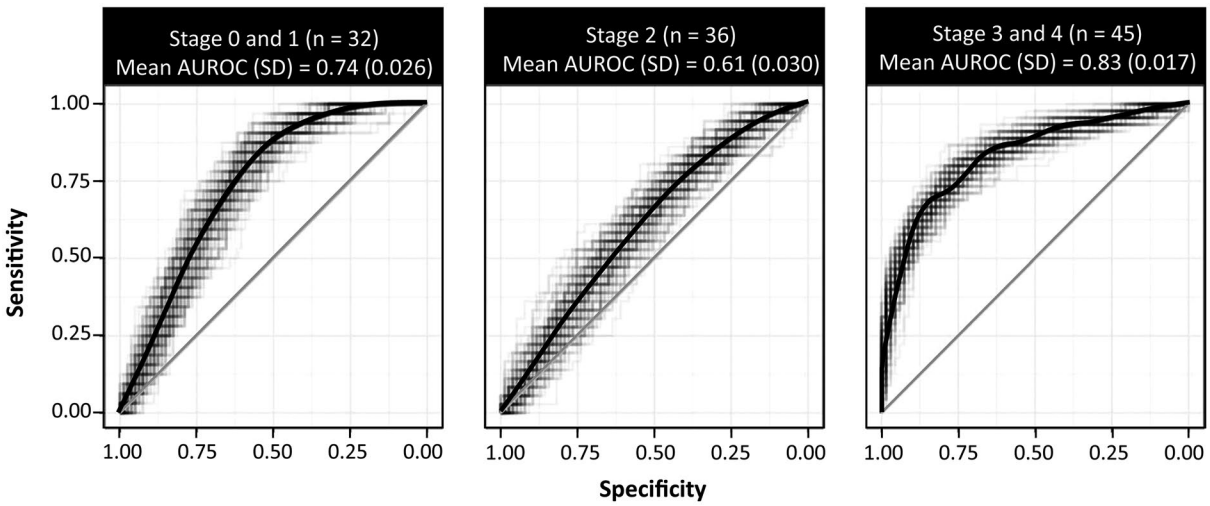

FIG. 6. Predictive models to discriminate fibrosis stages in patients with NASH. (A) Receiver operating characteristic (ROC) curve for the model to distinguish fibrosis stages 0-1 from stages 2-4. The model consists of SAP, fibrinogen, olfactomedin, and SHBG. (B) ROC curve for a model to classify patients with fibrosis stage 2. (C) ROC curve for the model to classify patients with advanced fibrosis. The model consists of LTBP4, IGF-1, VCAM1, IL1SRI, IL18Bpa, TSP2, collectin kidney 1, SHBG, TCCR, LIFsR, FBLN3, and PLXB2. Abbreviations: FBLN3, fibulin-3, IL1SRI, interleukin-1 soluble receptor type-1; LIFsR, leukemia inhibitory factor receptor, soluble; PLXB2, plexin-B2; TCCR, interleukin-27 receptor subunit alpha; TSP2, thrombospondin-2.

The analytical robustness of the SOMAscan platform has been previously established, and it has been used for proteomics-based studies in large cohorts.[20] The current study identified multiple proteins that are linked to biological processes that are related to cell stress/injury (i.e., GDF-15, FAS),[21,22] inflammation (i.e., collectin-11, IL18BP, C7, sCD163, TNF sR-II, SAP, VCAM1),[23-30] extracellular matrix remodeling (i.e., lumicans, TIMP, laminin),[31-33] and fibrogenesis (i.e., GAL3BP, LTBP4, TGFBI, FAP),[34-37] which contribute to chronic tissue injury and fibrosis. These biologically plausible results increase the likelihood that the findings are real and likely to be replicated in future studies. Importantly, these results further attest to the biological complexity of underlying disease progression. Finally, the clustering analysis opens the possibility of identifying clusters of individuals with specific patterns of biological pathway dysfunction that are reflected in the circulating proteome.

Furthermore, it is interesting to note the inverse relationship of IGF1 and its major binding protein IGFBP-3 with the severity of fibrosis. We have previously reported this relationship with NASH and have further shown that secretion of IGFBP-3, the most abundant IGF binding protein, from hepatic macrophages is decreased under states of lipotoxic stress with reversal of its normal inhibitory tone on hepatic IL-8 synthesis and secretion.[38] The current study confirms the inverse relationship with disease severity and extends it to advanced fibrosis stages (stages 3-4). These findings further support a strong biological basis of the identified proteins, their relationship to advanced fibrosis, and their use in a diagnostic panel to identify individuals with NASH and advanced fibrosis. Furthermore, other IGF binding proteins, such as IGFBP-1, IGFBP-2 and IGFBP-7, which regulate IGF tissue availability and activity, were higher in patients with advanced fibrosis. IGF-1 has been reported in preclinical models to inactivate stellate cells and reduce fibrosis in the liver,[39] whereas IGFBP-1 and IGFBP-7 have been reported to promote fibrosis in preclinical models.[40,41]

Hypogonadism has also been reported in patients with NAFLD[42] and IGF-1 deficiency may contribute to this phenotype. We also observed an elevation of SHBG in stage 3-4 patients and a reciprocal relationship of SHBG and IGF-1, consistent with a previous report[43], and of SHBG and the IGF-1:IGFBP-1 ratio. SHBG regulates free sex hormone levels, especially testosterone, and has been reported to be elevated in patients with NASH with cirrhosis.[44] These observations suggest that a reduction in the availability

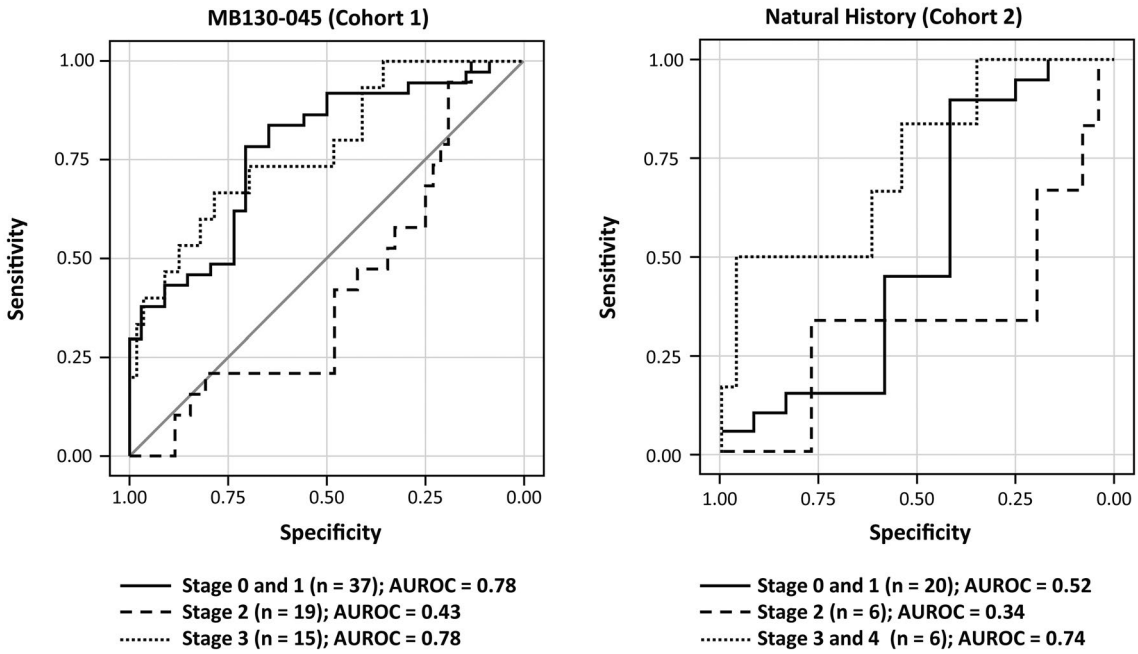

FIG. 7. Predictive performance of models in independent validation cohorts. ROC curves are shown for models for stage 0-1 (solid black line), stage 2 (dashed line), and stage 3-4 (dotted line).

of free testosterone may contribute to hypogonadism observed in patients with advanced fibrosis.[44,45] It should be noted that testosterone replacement therapy has been proposed to treat patients with NASH, such as in trial NCT01919294.

In addition to biological plausibility, the utility of a biomarker panel is determined by its diagnostic performance. The panel selected for identification of advanced fibrosis, consisting of proteins linked to diverse pathophysiological pathways, yielded an AUROC of 0.83, supporting the utility of this proteomics-based panel for this purpose. Similar performance in the validation cohorts reinforced the potential utility of this panel. Furthermore, for all three patient cohorts, the performance of the SOMAscan model for identifying those patients with advanced fibrosis was equal to, if not better than, that of several other models (FIB-4, APRI, and NFS), which have been previously used to distinguish patients with early versus advanced fibrosis. We attempted to select biomarkers based on serum proteomics data; however, the clinical applicability of the protein signature is currently limited, given the semi-quantitative nature of proteomics and the complexity of the algorithm. Further studies to characterize the selected biomarkers in patients at risk of NAFLD are warranted.

It was disappointing that a protein panel that could distinguish between stages 0-1 and ≥2 with high accuracy was not identified in this analysis. This could be a function of several factors, such as the population studied and the size of the derivation cohort, the accuracy of the biopsy-based diagnosis of stage 2 fibrosis, or that the nature of the stage 2 fibrosis may be actively progressive or regressive. Patients with stage 2 fibrosis in the discovery cohort were diverse at the molecular level, indicated by the clustering analysis that showed some patients with stage 2 fibrosis clustered with patients with stage 3-4 fibrosis, while others were clustered with patients with stage 0-1 fibrosis. It is also possible that a true signature distinguishing fibrosis stages 0–1 from stage ≥2 cannot be identified from circulating protein levels in NASH; however, given the ability to identify a proteomic signature for NASH with advanced fibrosis, we believe additional studies are needed to definitively determine whether a signature of clinically significant fibrosis (fibrosis stage ≥2) can be identified.

Analysis of the serum proteins also did not distinguish between patients with stage 3 versus stage 4 fibrosis, which is not entirely unexpected. The progression of fibrosis across the entire liver is

not synchronized, and a substantial proportion of patients with bridging fibrosis (stage 3) are found to have cirrhosis (stage 4) when additional biopsy cores are obtained.[46] This is corroborated by clinical trial experience, in which a subset of individuals with stage 3 fibrosis have been excluded from clinical trials because they have thrombocytopenia, a surrogate measure of portal hypertension, which typically develops when cirrhosis is present. Consistent with our clustering analysis (Fig. 2), some patients with stage 3 fibrosis were molecularly clustered with those with stage 4 fibrosis and vice versa. A number of proteins were significantly differentially expressed in S2 enriched with patients with stage 4 fibrosis compared to S1 enriched with those who had stage 3 fibrosis. The S2 patient subtype may better reflect the pathophysiology of cirrhosis than fibrosis staging. Further studies with a larger cohort of patients with stage 3-4 fibrosis are needed to further characterize the S2 and S1 subtypes.

The current study also provides pathophysiological insights into disease progression and the associated increase in risk of hepatocellular carcinoma. Notably, median levels of glypican-3, which is markedly elevated in patients with hepatocellular carcinoma, was 1.5-fold higher in patients with stage 4 versus stage 3 fibrosis (Supporting Table S5). Glypican-3 elevation in some patients with stage 4 fibrosis may suggest an early molecular event in hepatocarcinogenesis in patients with cirrhosis and may have biological significance in disease progression.[47,48] These differentially expressed proteins in patients with stage 4 fibrosis require further study to understand their roles in NASH disease progression.

A limitation of the current study is the overall size of the cohorts and the limited number of patients with stage 4 fibrosis. It should be noted that the validation cohorts lacked patients with stage 4 fibrosis, which may explain the slightly lower AUROC, although no protein met the FDR cutoff of 0.01 comparing fibrosis stage 3 with stage 4. Regardless, the reproducible performance of the panel to identify those individuals with advanced fibrosis supports its continued development. Another significant limitation is that all cohorts in this study consisted of patients with biopsy-proven NASH, which does not represent the spectrum of patients with NAFLD who are diagnosed in primary care or hepatology clinics. Thus, the identified biomarker panel is exploratory at present and must be further evaluated to determine its clinical applicability. In future studies, quantitative assays will need to be developed to further evaluate this biomarker panel; it will need to be evaluated in patient populations with risk factors for NAFLD or NASH in gastroenterology and hepatology clinics, as well as diabetes clinics and the primary care setting, in order to meet regulatory standards for biomarker qualification. The results from this exploratory analysis strongly support future studies.

In summary, an unbiased proteomics serum analysis identified a panel of biomarkers reflecting disease processes known to be involved in the pathophysiological cascade of NASH such as cell injury, inflammation, fibrogenesis, and tissue remodeling that were associated with advanced fibrosis. These results demonstrate that this exploratory biomarker panel warrants further evaluation as a diagnostic tool for advanced fibrosis in patients with NASH and provides rationale for confirmation of these findings in additional large, well-characterized, intended-use populations.

Acknowledgment: The authors thank Michael Idowu, M.D., and Melissa J. Contos, M.D., of Virginia Commonwealth University for conducting the histological assessments. The authors also thank the anonymous reviewers for their insightful feedback. Medical writing and editorial support was provided by Apurva Davé, Ph.D., and Amanda Martin, Ph.D., of Medical Expressions (Chicago, IL), and was funded by Bristol Myers Squibb.

REFERENCES

1) Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the study of Liver Diseases. Hepatology 2018;67:328-357.
2) Kim D, Li AA, Perumpail BJ, Gadiparthi C, Kim W, Cholankeril G, et al. Changing trends in etiology-based and ethnicity-based annual mortality rates of cirrhosis and hepatocellular carcinoma in the United States. Hepatology 2019;69:1064-1074.
3) Younossi Z, Stepanova M, Ong JP, Jacobson IM, Bugianesi E, Duseja A, et al. Nonalcoholic steatohepatitis is the fastest growing cause of hepatocellular carcinoma in liver transplant candidates. Clin Gastroenterol Hepatol 2019;17:748-755.
4) Younossi ZM. Non-alcoholic fatty liver disease—a global public health perspective. J Hepatol 2019;70:531-544.
5) Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology 2015;149:389-397.
6) Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific

mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;61:1547-1554.
7) Sanyal AJ, Harrison SA, Ratziu V, Abdelmalek MF, Diehl AM, Caldwell S, et al. The natural history of advanced fibrosis due to nonalcoholic steatohepatitis: data from the simtuzumab trials. Hepatology 2019;70:1913-1927.
8) Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med 2018;24:908-922.
9) Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. J Hepatol 2018;68:305-315.
10) Dulai PS, Sirlin CB, Loomba R. MRI and MRE for non-invasive quantitative assessment of hepatic steatosis and fibrosis in NAFLD and NASH: clinical trials to clinical practice. J Hepatol 2016;65:1006-1016.
11) Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody E, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 2010;5:e15004.
12) Begic E, Hadzidedic S, Kulaglic A, Ramic-Brkic B, Begic Z, Causevic M. SOMAscan-based proteomic measurements of plasma brain natriuretic peptide are decreased in mild cognitive impairment and in Alzheimer's dementia patients. PLoS One 2019;14:e0212261.
13) Ko D, Benson MD, Ngo D, Yang Q, Larson MG, Wang TJ, et al. Proteomics profiling and risk of new-onset atrial fibrillation: Framingham Heart Study. J Am Heart Assoc 2019;8:e010976.
14) Wood GC, Chu X, Argyropoulos G, Benotti P, Rolston D, Mirshahi T, et al. A multi-component classifier for nonalcoholic fatty liver disease (NAFLD) based on genomic, proteomic, and phenomic data domains. Sci Rep 2017;7:43238.
15) Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313-1321.
16) Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption–II. Addiction 1993;88:791-804.
17) Sanyal A, Charles ED, Neuschwander-Tetri BA, Loomba R, Harrison SA, Abdelmalek MF, et al. Pegbelfermin (BMS-986036), a PEGylated fibroblast growth factor 21 analogue, in patients with non-alcoholic steatohepatitis: a randomised, double-blind, placebo-controlled, phase 2a trial. Lancet 2019;392:2705-2717.
18) Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, et al. Assessment of variability in the SOMAscan assay. Sci Rep 2017;7:14248.
19) Ritchie ME, Phipson B, Wu DI, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
20) Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, et al. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. JAMA 2016;315:2532-2541.
21) Feldstein AE, Canbay A, Angulo P, Taniai M, Burgart LJ, Lindor KD, et al. Hepatocyte apoptosis and Fas expression are prominent features of human nonalcoholic steatohepatitis. Gastroenterology 2003;125:437-443.
22) Hsiao EC, Koniaris LG, Zimmers-Koniaris T, Sebald SM, Huynh TV, Lee SJ. Characterization of growth-differentiation factor 15, a transforming growth factor beta superfamily member induced following liver injury. Mol Cell Biol 2000;20:3742-3751.
23) Bottazzi B, Inforzato A, Messa M, Barbagallo M, Magrini E, Garlanda C, et al. The pentraxins PTX3 and SAP in innate immunity, regulation of inflammation and tissue remodelling. J Hepatol 2016;64:1416-1427.
24) Hayder H, Blanden RV, Korner H, Riminton DS, Sedgwick JD, Mullbacher A. Adenovirus-induced liver pathology is mediated through TNF receptors I and II but is independent of TNF or lymphotoxin. J Immunol 1999;163:1516-1520.
25) Lefere S, Van de Velde F, Devisscher L, Bekaert M, Raevens S, Verhelst X, et al. Serum vascular cell adhesion molecule-1 predicts significant liver fibrosis in non-alcoholic fatty liver disease. Int J Obes (Lond) 2017;41:1207-1213.
26) Lidofsky A, Holmes JA, Feeney ER, Kruger AJ, Salloum S, Zheng H, et al. Macrophage activation marker soluble CD163 is a dynamic marker of liver fibrogenesis in human immunodeficiency virus/hepatitis C virus coinfection. J Infect Dis 2018;218:1394-1403.
27) Lin CJ, Hu ZG, Yuan GD, Lei B, He SQ. Complements are involved in alcoholic fatty liver disease, hepatitis and fibrosis. World J Hepatol 2018;10:662-669.
28) Spahr L, Garcia I, Bresson-Hadni S, Rubbia-Brandt L, Guler R, Olleros M, et al. Circulating concentrations of interleukin-18, interleukin-18 binding protein, and gamma interferon in patients with alcoholic hepatitis. Liver Int 2004;24:582-587.
29) Wu W, Liu C, Farrar CA, Ma L, Dong X, Sacks SH, et al. Collectin-11 promotes the development of renal tubulointerstitial fibrosis. J Am Soc Nephrol 2018;29:168-181.
30) Zhang M, Ye Y, Wang F, Zhu J, Zhao Q, Zheng Y, et al. Liver myofibroblasts up-regulate monocyte CD163 expression via PGE2 during hepatitis B induced liver failure. J Transl Med 2014;12:60.
31) Clément B, Rescan P-Y, Baffet G, Loréal O, Lehry D, Campion J-P, et al. Hepatocytes may produce laminin in fibrotic liver and in primary culture. Hepatology 1988;8:794-803.
32) Krishnan A, Li X, Kao W-Y, Viker K, Butters K, Masuoka H, et al. Lumican, an extracellular matrix proteoglycan, is a novel requisite for hepatic fibrosis. Lab Invest 2012;92:1712-1725.
33) Munsterman ID, Kendall TJ, Khelil N, Popa M, Lomme R, Drenth JPH, et al. Extracellular matrix components indicate remodelling activity in different fibrosis stages of human non-alcoholic fatty liver disease. Histopathology 2018;73:612-621.
34) Kamada Y, Ono M, Hyogo H, Fujii H, Sumida Y, Yamada M, et al. Use of Mac-2 binding protein as a biomarker for nonalcoholic fatty liver disease diagnosis. Hepatol Commun 2017;1:780-791.
35) LeBaron RG, Bezverkov KI, Zimber MP, Pavelec R, Skonier J, Purchio AF. Beta IG-H3, a novel secretory protein inducible by transforming growth factor-beta, is present in normal skin and promotes the adhesion and spreading of dermal fibroblasts in vitro. J Invest Dermatol 1995;104:844-849.
36) Lu J, Liu Q, Wang L, Tu W, Chu H, Ding W, et al. Increased expression of latent TGF-beta-binding protein 4 affects the fibrotic process in scleroderma by TGF-beta/SMAD signaling. Lab Invest 2017;97:591-601.
37) Wang XM, Yu DM, McCaughan GW, Gorrell MD. Fibroblast activation protein increases apoptosis, cell adhesion, and migration by the LX-2 human stellate cell line. Hepatology 2005;42:935-945.
38) Min H-K, Maruyama H, Jang BK, Shimada M, Mirshahi F, Ren S, et al. Suppression of IGF binding protein-3 by palmitate promotes hepatic inflammatory responses. FASEB J 2016;30:4071-4082.
39) Sanz S, Pucilowska JB, Liu S, Rodriguez-Ortigosa CM, Lund PK, Brenner DA, et al. Expression of insulin-like growth factor I by activated hepatic stellate cells reduces fibrogenesis and enhances regeneration after liver injury. Gut 2005;54:134-141.
40) Colak Y, Senates E, Ozturk O, Yilmaz Y, Zemheri E, Yilmaz Enc F, et al. Serum concentrations of human insulin-like growth

factor-1 and levels of insulin-like growth factor-binding protein-5 in patients with nonalcoholic fatty liver disease: association with liver histology. Eur J Gastroenterol Hepatol 2012;24:255-261.
41) Hagstrom H, Stal P, Hultcrantz R, Brismar K, Ansurudeen I. IGFBP-1 and IGF-I as markers for advanced fibrosis in NAFLD - a pilot study. Scand J Gastroenterol 2017;52:1427-1434.
42) Hazlehurst JM, Tomlinson JW. Non-alcoholic fatty liver disease in common endocrine disorders. Eur J Endocrinol 2013;169:R27-R37.
43) Kaaks R, Lukanova A, Rinaldi S, Biessy C, Söderberg S, Olsson T, et al. Interrelationships between plasma testosterone, SHBG, IGF-I, insulin and leptin in prostate cancer cases and controls. Eur J Cancer Prev 2003;12:309-315.
44) Handelsman DJ, Strasser S, McDonald JA, Conway AJ, McCaughan GW. Hypothalamic-pituitary-testicular function in end-stage non-alcoholic liver disease before and after liver transplantation. Clin Endocrinol (Oxf) 1995;43:331-337.
45) Sinclair M, Grossmann M, Gow PJ, Angus PW. Testosterone in men with advanced liver disease: abnormalities and implications. J Gastroenterol Hepatol 2015;30:244-251.
46) Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898-1906.
47) Liu X, Wang SK, Zhang K, Zhang H, Pan Q, Liu Z, et al. Expression of glypican 3 enriches hepatocellular carcinoma development-related genes and associates with carcinogenesis in cirrhotic livers. Carcinogenesis 2015;36:232-242.
48) Montalbano M, Rastellini C, McGuire JT, Prajapati J, Shirafkan A, Vento R, et al. Role of glypican-3 in the growth, migration and invasion of primary hepatocytes isolated from patients with hepatocellular carcinoma. Cell Oncol (Dordr) 2018;41: 169-184.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1670/suppinfo.