Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis

W.-M. WANG1, W.-S. ZHANG2, Z.-G. YANG2

1Department of Traditional Chinese Medicine, Beijing Tian Tan Hospital, Capital Medical University, Beijing, China
2Department of Integrative Medicine, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

Abstract. – OBJECTIVE: The crosstalk between Toll-like receptor 4 (TLR-4) and lipopolysaccharide (LPS) accounts for liver fibrosis progression. This study aimed to investigate the predictive performance of altered genes induced by TLR-4 and LPS challenge for advanced liver fibrosis.

MATERIALS AND METHODS: The overlapping differentially expressed genes (DEGs) of TLR-4 and LPS challenge models from the Gene Expression Omnibus (GEO) database were screened and included in the random forest analysis to identify potential candidates for predicting advanced liver fibrosis in the GSE84044 dataset. The roles of the identified candidates in liver injury development and activation of hepatic stellate cells (HSCs) were also addressed.

RESULTS: Among the overlapping DEGs in the GSE30485, GSE33446 and GSE166488 datasets, vimentin (VIM) was the most important gene for predicting advanced liver fibrosis (S ≥ 2) by the random forest model. In the GSE84044 dataset, VIM was positively correlated with liver fibrosis (r = 0.68, 95% CI = 0.57-0.76, p < 0.0001), and accurately predicted advanced liver fibrosis (AUC = 0.85, 95% CI = 0.78-0.91), both in males (AUC = 0.84, 95% CI = 0.76-0.92) and females (AUC = 0.87, 95% CI = 0.76-0.99). VIM was significantly upregulated in various liver diseases (cirrhosis, liver failure, chronic hepatitis B and fatty liver disease) and liver injury models (ANIT, BDL, CCI, and DMN). Additionally, VIM was correlated with HSC regulators (TGFβ, PDGF, CTGF and BMP7) and overexpressed in activated HSCs (p < 0.05). Enrichment analysis indicated that VIM-induced gene alterations were involved in the cytosolic DNA-sensing pathway, Toll-like receptor signaling pathway, etc.

CONCLUSIONS: VIM could predict advanced liver fibrosis in CHB patients and is mainly involved in the activation of HSCs and profibrotic signaling pathways.

Key Words: TLR-4, Lipopolysaccharide, VIM, Liver fibrosis, Cirrhosis, Hepatic stellate cells, TGFβ.

Introduction

Liver fibrosis is the precursor to cirrhosis and liver cancer, which finally progress to liver failure and death1,2. A variety of chronic liver injuries, including virological, alcoholic and nonalcoholic, autoimmune, metabolic, and chemical-toxic liver diseases can lead to liver fibrosis3. Hepatic stellate cell (HSC) activation and excessive deposition of extracellular matrix (ECM) are the major pathophysiological changes during the development of liver fibrosis4,6. In addition, various signaling pathways are involved in liver fibrosis progression, such as transforming growth factor β (TGFβ) and nuclear factor-κB (NF-κB)7,8. Beyond this, intestinal bacterial microflora and a functional Toll-like receptor 4 (TLR4) are required for hepatic fibrogenesis, which can enhance TGFβ and NF-κB signaling pathways9,11.

The TLR4/myeloid differentiation-2 (MD-2) heterodimer is a recognition molecule of lipopolysaccharide (LPS), which is a membrane component of Gram-negative bacteria12,13. The activated TLR4/MD-2/LPS complex triggers intracellular pathways resulting in the activation of nuclear transcription factors and promoting the production and accumulation of proinflammatory cytokines14,15. In wild type mice, LPS could markedly induce serum liver dysfunction and liver pathological changes, increase the levels of inflammatory cytokines, and upregulate TLR4 signaling cascade markers. Interestingly, these pathological changes are greatly alleviated in TLR4-knockout mice. LPS stimulation provokes the pathological responses in primary Kupffer cells isolated from wild type and TLR4-knockout mice16,18. Considering the crosstalk between LPS and TLR4, we assume that gene expression changes induced by LPS challenge and TLR4-mutant toxic liver inju-
ries might contribute to the liver fibrosis progression and serve as predictive candidates.

Since liver fibrosis is the result of various chronic liver injuries and the precursor to end-stage liver damage, early diagnosis and assessment are critical for the reversal of this pathological stage. In the present study, we aim to elucidate the gene alterations induced by LPS challenge and TLR4-mutant toxic liver injuries and investigate the correlations between these alterations and liver fibrosis progression, hoping that these data could provide novel diagnostic candidates, as well as useful insights into the pathogenesis and mechanisms of liver fibrosis.

**Materials and Methods**

**Ethics Statement**

All participants provided written informed consent. The protocol of the primary study was approved by the Ethics Committee of Ruijin Hospital, Shanghai Jiaotong University, School of Medicine. The protocol of this secondary analysis was reviewed and approved by the Ethics Committee, Shanghai Public Health Clinical Center, Fudan University.

**Patients**

The GSE84044 dataset in the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database was used to screen candidates for predicting advanced liver fibrosis. A total of 124 chronic hepatitis B (CHB) patients in this dataset who underwent liver biopsy were included. The age, sex, and histopathology stages of these participants were publicly available. As stated by Wang et al., the diagnosis of CHB patients was determined according to the criteria from the Asian Pacific Association for the Study of the Liver (APASL). Patients with the following conditions were excluded from the study: 1) Use of any antiviral therapies or immunosuppressive drugs within six months before sampling; 2) Infection with human immunodeficiency virus (HIV) or with a hepatitis virus other than hepatitis B virus (HBV), autoimmune liver disease, drug induced liver injury, alcoholic liver disease or hepatocellular carcinoma.

**Outcome**

The primary outcome was advanced liver fibrosis, which was defined as histological fibrosis staging $S \geq 2$. The histopathological diagnosis of all the liver biopsies of CHB patients was conducted by two experienced pathologists from the Pathology Department of Shanghai Fudan University, School of Medicine. The fibrosis staging (Scheuer S) and inflammation grading (Scheuer G) were calculated according to the Scheuer scoring system, namely, $S$ 0-4 and $G$ 0-4.

**Differentially Expressed Genes (DEGs) Identification**

This identification framework was applied to all GEO series included in this study. The affy package was used to normalize the microarray data of raw.cel files from each GEO dataset with the quantile method of robust multichip analysis (RMA). The limma package was used to compare gene expression. Missing gene expression data were imputed with the k-nearest neighbor method by impute index. The average gene expression was calculated when multiple probes of the genes existed. All platforms and samples of each microarray series were downloaded from the GEO database. DEGs were identified with the criterion of a $|\log 2 \text{FC}| > 1.0$ and adjusted $p$-value < 0.05.

The keywords “Toll-like receptor”, “TLR4”, “lipopolysaccharide”, and “LPS” were used to search potential profiles for subsequent analysis in the GEO database. GSE33446, GSE30485, and GSE166488 were included in the analysis for the identification of overlapping DEGs. The GSE33446 and GSE30485 profiles compared the gene expression in the mouse liver induced by carbon tetrachloride ($\text{CCl}_4$) and diethylnitrosamine (DEN) in the TLR4 mutant models, and GSE166488 compared the gene expression in the mouse liver induced by LPS injection. The details of these profiles are presented in the Supplemental Materials.

**Enrichment Analysis**

The Investigate Gene Sets tool of the Molecular Signatures Database (MSigDB) version 7.5.1 in the Gene Set Enrichment Analysis (GSEA, http://www.gsea-msigdb.org/) was used for functional enrichment of the DEGs. Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome pathway analyses and Gene Ontology (GO) analyses were performed in the GSEA database. The enrichment terms with a false discovery rate (FDR) less than 0.05 were included. Human or mouse species were determined by the sample type.
**Random Forest Model Establishment**

The randomForest package was used to screen potential candidates for predicting advanced liver fibrosis. The 124 CHB patients in the GSE84044 dataset were randomly divided into training and validation cohorts in a 5:5 ratio using the "caret" package. In the process of constructing the random forest model, the variable importance of the output results of the Gini coefficient method was calculated from the perspective of decreasing accuracy and decreasing mean square error. DEGs with an importance greater than 3 were identified as candidate genes for subsequent analysis.

**Statistical Analysis**

Differences in gene expression between the individual groups were analyzed using Student’s t-test, the Mann-Whitney test or two-way ANOVA based on variable types by GraphPad Prism 8 (GraphPad Software, San Diego, CA, USA). The OptimalCutpoints package in the R program was used to perform ROC analysis to evaluate the predictive values of potential candidates for the liver fibrosis stage. Correlation analysis was addressed by Spearman or Pearson methods. Stata software version 16.1 (Stata Corp LLC, TX, USA) was used for logistic regression and correlation analysis. A two-tailed p < 0.05 was considered significant.

**Results**

**Overlapping DEGs of TLR4-Mutant Toxic and LPS-Induced Liver Injury Models**

A total of 49 overlapping DEGs were identified among GSE33446, GSE30485, and GSE166485 (Figure 1A). As summarized in Figure 1B, enrichment analysis indicated that these DEGs were mainly involved in biological processes (BP) including defense response, programmed cell death, and regulation of apoptosis. These DEGs had the molecular function (MF) in Toll-like receptor 4 binding. Additionally, the cellular component (CC) of these overlapping DEGs was collagen containing extracellular matrix. Reactome analysis showed that these overlapping DEGs were significantly enriched in the innate immune system, neutrophil degranulation, etc. (Figure 1B). Since 15 DEGs (Al426330, AW112010, Ces1e, Gm2788, Hsd3b5, Ifi271a2, Ly6a, Mup5, Mup10, Slco1a1, Slfn4, Ugt2b1, 4921539H07Rik, 4933438K21Rik, and 9130409J23Rik) were not expressed in human species, the levels of CD163, CD52, CTSC, DYN-L1, ECSCR, EGR1, EVI2A, GPCPD1, ICAM1, IFI44, LCN2, LGALS3, ME1, MOXD1, NCF4, NUCB2, PBK, PROCR, S100A9, and VIM were significantly upregulated in CHB patients with liver fibrosis S stage ≥ 2 (p < 0.05, Figure 1C), while APOA4 and NUPR1 were significantly downregulated in this population (p < 0.001, Figure 1C).

**VIM Expression in Liver Diseases and Liver Injury Models**

As shown in Figure 4, VIM was significantly upregulated in cirrhosis patients compared to normal controls (p < 0.05, GSE7741, GSE14323 and GSE103580, Figure 4A). The levels of VIM were significantly higher in patients with...
Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis

HBV-related liver failure than in healthy individuals \((p < 0.001, \text{GSE}14668^{38}, \text{GSE}38941^{39} \text{and GSE}96851^{40}, \text{Figure 4B})\) and significantly elevated in HBV patients in the immune clearance stage compared to inactive carriers or immune tolerant patients \((p < 0.05, \text{GSE}65359, \text{Figure 4B})\). In addition, interferon therapy significantly decreased VIM expression in CHB patients \((p < 0.001, \text{GSE}66698^{41}, \text{Figure 4B})\). Similarly, VIM was significantly overexpressed in patients suffering from NAFLD or NASH compared to normal controls \((p < 0.05, \text{GSE}35961^{42}, \text{GSE}59045^{43} \text{and GSE}63067^{44}, \text{Figure 4C})\), and it was also significantly upregulated in the high-fat-diet model \((p < 0.05, \text{GSE}57425^{45} \text{and GSE}23740, \text{Figure 4C})\).

In addition, VIM was significantly upregulated in liver injury models, namely, ANIT \((p < 0.01, \text{GSE}122184^{46} \text{and GSE}72387^{47}, \text{Figure 4D})\), BDL \((p < 0.0001, \text{GSE}15249^{48}, \text{Figure 4E})\), CCl\(_4\) \((p < 0.05, \text{GSE}167033^{49}, \text{GSE}122184^{46} \text{and GSE}89147^{50}, \text{Figure 4F})\), and DMN \((p < 0.001, \text{GSE}122184^{46}, \text{GSE}68110^{51}, \text{GSE}58032^{52} \text{and GSE}44783^{53}, \text{Figure 4G})\) compared to normal livers. The levels of VIM were also significantly overexpressed in LPS-induced livers compared to normal controls in the GSE37546 \(^{54} \text{and GSE}55084^{55} \text{datasets} \((p < 0.05, \text{Figure 4H})\).

**Correlation Between VIM and HSCs**

Since the activation of HSCs is the core process of liver fibrosis, the VIM levels in HSCs were in-
investigated in four GEO series. As shown in Figure 4A, VIM was significantly upregulated in growing HSCs compared to senescent HSCs \( (p < 0.05, \text{GSE119545}^6, \text{Figure 5A}) \), myofibroblasts transited HSCs compared to quiescent HSCs \( (p < 0.0001, \text{GSE349495}^7, \text{Figure 5A}) \), HSCs compared to liver stem/progenitor cells, hepatocytes, and liver sinusoidal endothelial cells \( (p < 0.01, \text{GSE49995}^8 \text{and GSE68000}^9, \text{Figure 5A}) \), and activated HSCs compared to quiescent HSCs and reverted HSCs \( (p < 0.05, \text{GSE68000}^9 \text{and GSE68001}^{10}, \text{Figure 5A}) \).

TGFβ, PDGF, and CTGF are the main activators of HSCs, and BMP7 is an inhibitor of HSCs. The correlations between VIM and these HSC regulators in the GSE84044 dataset are presented in Figure 5B. The HSC regulators with correlation coefficient \(| r | \geq 0.30\) are shown in Table 1.

### Table 1. Baseline characteristics of CHB patients included in this study.

<table>
<thead>
<tr>
<th>Variables</th>
<th>( S &lt; 2 ) (n = 62)</th>
<th>( S \geq 2 ) (n = 62)</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (Interquartile range, IQR)</td>
<td>35.5 (31, 46)</td>
<td>43.5 (37, 55)</td>
<td>0.001</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>48 (77.4)</td>
<td>40 (64.5)</td>
<td>0.114</td>
</tr>
<tr>
<td>G grade, n (%)</td>
<td></td>
<td></td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>0</td>
<td>30 (48.4)</td>
<td>7 (11.3)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21 (33.9)</td>
<td>12 (19.4)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10 (16.1)</td>
<td>24 (38.7)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1 (1.6)</td>
<td>14 (22.6)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0 (0)</td>
<td>5 (8.1)</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.** The establishment of random forest. The error of random forest trees (A), and the variable importance of the output results of the Gini coefficient method (B).
Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis

0.5 is detailed in Figure 5C. VIM was positively correlated with TGFβ2, PDGFA, PDGFD, and CTGF ($p < 0.0001$, Figure 5C) and negatively correlated with BMP7 ($p < 0.0001$, Figure 5C).

**Signaling Pathways and GO Enrichment of VIM-Induced DEGs**

We searched the GEO database with the keywords “VIM” and “vimentin”, and the GSE63653
Figure 4. The VIM expression in cirrhosis (A), HBV infected patients (B), fatty liver diseases (C), ANIT (D), BDL (E), CCl4 (F), DMN (G), and LPS-induced liver injury (H) models.
Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis

Figure 5. The VIM expression in hepatis stellate cells (A); the correlation coefficient between VIM and HSC regulators (B); the statistics of markers correlated with VIM with coefficient $|r| \geq 0.5$ (C).
VIM is expressed in mesenchymal cells with fibrosis, but not frequently in hepatic fibrosis. Mice are widely used to explore the process of liver fibrosis. VIM is predominantly expressed in HSCs and is also present in vascular smooth muscle cells and portal fibroblasts. The expression of VIM in liver fibrosis is still controversial. Immunohistochemistry identified that VIM was positive in fibrotic areas in human liver fibrosis patients. Endogenous VIM was remarkably expressed in the CCl4 and BDL-induced liver fibrosis models, which is consistent with our results. In another report, VIM was upregulated in nonorthotopic liver transplantation (OLT) and post-LOT patients compared to nonfibrosis populations, but did not reach statistical significance. In CCl4-induced liver fibrosis, VIM was not observed in the hepatocytes, even upon stimulation with TGFβ1. Our results may further explain the role of VIM in the progression of liver fibrosis. Given the previous reports and our results, we believe that VIM should be a promising predictor for advanced liver fibrosis in clinical practice.

Our results showed that VIM is involved in the activation of HSCs and might be linked to profibrotic signaling pathways, including TGFβ, TLR4, and NF-κB. VIM regulates the assembly of focal adhesions through collagen and affects signaling pathways that control extracellular matrix remodeling. A previous report revealed that CCl4 and dextran sulfate sodium treatment could increase the expression of VIM, TGFβ, TLR4, and NF-κB p65, leading to activation of HSCs and the TLR4 signaling pathway. In addition, VIM mRNA in primary hepatocytes was downregulated in matrix metalloproteinase-19 (MMP-19) knockout mice, which showed a decreased response to TGFβ1 stimulation. Unfortunately, we have not investigated the causality of VIM upregulation, profibrotic pathways, and HSC activation through experimental assays. Future research should focus on the clinical application and basic mechanisms of VIM in the development of liver fibrosis.

Some limitations existed in this study. First, this was a secondary analysis from a public database with inadequate clinical characteristics; for example, liver function tests, serum liver fibrosis parameters and virological markers were not obtained and adjusted, which may have resulted in biases in the predictive power of VIM for advanced liver fibrosis. Second, the serum VIM levels in liver fibrosis patients were not addressed, which limits the application of this candidate in clinical practice in the current situation. Third, the mechanisms of VIM in the progression of liver fi-
Figure 6. The DEGs induced by VIM knockout in macrophages in the GSE63653 (A), and the KEGG (B), Reactome (C) and GO enrichment (D) of these DEGs.
Fibrosis were discussed superficially in this study, and no experimental clarifications were conducted in vivo and in vitro. Fourth, VIM might be a novel diagnostic candidate for liver fibrosis, but research on VIM as a therapeutic target still needs to go further. Considering the complicated mechanisms of liver fibrosis, genetic and epigenetic aspects of VIM in liver fibrosis progression should be investigated in depth.

Conclusions

In this study, we identified VIM as a promising predictor for advanced liver fibrosis in CHB patients. Additionally, VIM was widely upregulated in various liver injuries and was involved in the activation of HSCs and HSC regulators. Moreover, VIM might be linked to profibrotic signaling pathways, including TGFβ, TLR4, and NF-κB. Considering the current evidence, VIM should be used as a novel diagnostic candidate for liver fibrosis. More prospective studies should be conducted to evaluate the diagnostic values of serum VIM for liver fibrosis.

Acknowledgements

This work was supported by fundings through the Shanghai Science and Technology Committee [grant number 19401931600] and the Shanghai Municipal Health Commission [grant number 2020LZ001]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors’ Contributions

Z.Y. and W.Z. conceived and designed the study. W.W. wrote the manuscript. Z.Y., and W.Z. rewriting the manuscript. W.W., W.Z., and Z.Y. analyzed and interpreted the data. All authors read and approved the final manuscript.

Conflicts of Interest

The authors declared no conflicts of interest in this work.

Data Availability

Datasets of the current study are available from the NCBI Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). All the datasets were available from Dr. Zongguo Yang upon reasonable request.

References

Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis


Vimentin (VIM) predicts advanced liver fibrosis in chronic hepatitis B patients: A random forest-derived analysis


75) Forbes SJ, Russo FP, Rey V, Burra P, Rugge M, Wright NA, Alison MR. A significant proportion of myofibroblasts are of bone marrow origin in human liver fibrosis. Gastroenterology 2004; 126: 955-963.


