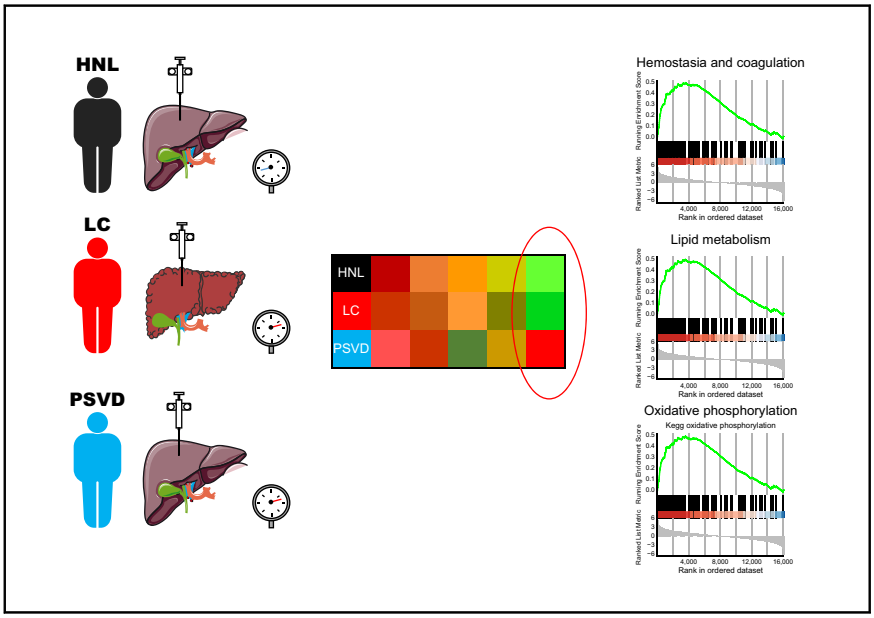
Research Article
Innovative Diagnostics, Modelling and Digital Hepatology

JOURNAL OF HEPATOLOGY

Co-expression gene network analysis reveals novel regulatory pathways involved in porto-sinusoidal vascular disease

Graphical abstract



Highlights

- Patients with PSVD have a unique transcriptomic profile.
- Pathways related to vascular homeostasis might be involved in PSVD development.
- Intravascular endothelial dysfunction is key in PSVD pathophysiology.
- Maintaining a healthy liver endothelium emerges as a promising therapeutic strategy.

Authors

Virginia Hernández-Gea, Genís Campreciós, Fabián Betancourt, …, Manuel Romero-Gómez, Juan-José Lozano, Juan Carlos García-Pagán

Correspondence

vihernandez@clinic.cat (V. Hernández-Gea), jcgarcia@clinic.cat (J.C. García-Pagán).

Lay summary

Porto-sinusoidal vascular disease is a rare but life-shortening disease that affects mainly young people. Knowledge of the disrupted pathways involved in its development will help to identify novel therapeutic targets and new treatments. Using a systems biology approach, we identify that pathways regulating endothelial function and tone may act as drivers of porto-sinusoidal vascular disease.

https://doi.org/10.1016/j.jhep.2021.05.014
 J. Hepatol. 2021, 75, 924–934

Research Article
Innovative Diagnostics, Modelling and Digital Hepatology

JOURNAL OF HEPATOLOGY

Co-expression gene network analysis reveals novel regulatory pathways involved in porto-sinusoidal vascular disease

Virginia Hernández-Gea[1,2,*,†], Genís Campreciós[1,2,†], Fabián Betancourt[1], Valeria Pérez-Campuzano[1], Susana Seijo[1], Alba Díaz[3], Rocío Gallego-Durán[4,2], Pol Olivas[1], Lara Orts[1], Marta Magaz[1], Anna Baiges[1,2], Fanny Turon[1,2], Julia Sidorova[5], Manuel Romero-Gómez[4,2], Juan-José Lozano[5], Juan Carlos García-Pagán[1,2,*]

[1]*Barcelona Hepatic Hemodynamic Laboratory, Liver Unit, Hospital Clínic, IDIBAPS, University of Barcelona, Health Care Provider of the European Reference Network on Rare Liver Disorders (ERN-Liver), Barcelona, Catalonia, Spain;* [2]*Centro de Investigación Biomédica Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Spain;* [3]*Pathology Department, Biomedical Diagnostic Centre, Hospital Clínic Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Catalonia, Spain;* [4]*Digestive Diseases Unit, Hospital Universitario Virgen del Rocío. SeLiver group, Instituto de Biomedicina de Sevilla/CSIC/Universidad de Sevilla, Sevilla, Spain;* [5]*Bioinformatic Platform, Centro de Investigación Biomédica Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Spain*

Background & Aims: Porto-sinusoidal vascular disease (PSVD) is a rare vascular liver disease of unknown etiology that causes portal hypertension. It usually affects young individuals and shortens live expectancy. The deregulated pathways involved in PSVD development are unknown and therefore we lack curative treatments. The purpose of this study was to integrate transcriptomic and clinical data by comprehensive network-based modeling in order to uncover altered biological processes in patients with PSVD.

Methods: We obtained liver tissue samples from 20 consecutive patients with PSVD and 21 sex- and age-matched patients with cirrhosis and 13 histologically normal livers (HNL) (initial cohort) and performed transcriptomic analysis. Microarray data were analyzed using weighted gene correlation network analysis to identify clusters of highly correlated genes differently expressed in patients with PSVD. We next evaluated the molecular pathways enriched in patients with PSVD and the core-related genes from the most significantly enriched pathways in patients with PSVD. Our main findings were validated using RNA sequencing in a different cohort of PSVD, cirrhosis and HNL (n = 8 for each group).

Results: Patients with PSVD have a distinctive genetic profile enriched mainly in canonical pathways involving hemostasis and coagulation but also lipid metabolism and oxidative phosphorylation. Serpin family (*SERPINC1*), the apolipoproteins (*APOA*, *APOB*, *APOC*), ATP synthases (*ATP5G1*, *ATP5B*), fibrinogen genes (*FGB*, *FGA*) and alpha-2-macroglobulin were identified as highly connective genes that may have an important role in PSVD pathogenesis.

Conclusion: PSVD has a unique transcriptomic profile and we have identified deregulation of pathways involved in vascular homeostasis as the main pathogenic event of disease development.

Lay summary: Porto-sinusoidal vascular disease is a rare but life-shortening disease that affects mainly young people. Knowledge of the disrupted pathways involved in its development will help to identify novel therapeutic targets and new treatments. Using a systems biology approach, we identify that pathways regulating endothelial function and tone may act as drivers of porto-sinusoidal vascular disease.

Introduction

Porto-sinusoidal vascular disease (PSVD) is a rare vascular liver disease of unknown etiology that causes portal hypertension. PSVD usually affects young individuals and progression of the disease leads to development of portal hypertension, which causes disability, and poor outcome. Many features of PSVD remain unknown and treatment is limited to management of portal hypertension complications. Its diagnosis is often challenging and relies on the combination of clinical and pathological criteria as we still lack a positive diagnostic test; it is sometimes misdiagnosed as early-stage cirrhosis. Moreover, histological findings may be very subtle and may be missed in the absence of a high clinical suspicion and a trained pathologist.[1]

The pathophysiological mechanisms responsible for PSVD development remain uncovered and hamper the development of treatments able to modify the natural history of the disease. A better insight into the biological pathways involved in PSVD development is essential to reveal disease drivers and identify new therapeutic targets and curative treatments.

Biological networks represent a valuable tool for uncovering novel deregulated pathways underlying molecular pathology. Contrary to high-throughput technology such as gene expression profiling, which analyzes genes based on differential expression,

Keywords: porto sinusoidal vascular disease; non-cirrhotic portal hypertension; portal hypertension; vascular liver disease; trasncriptomic analysis; system biology.
Received 27 September 2020; received in revised form 14 May 2021; accepted 17 May 2021; available online 28 May 2021

* Corresponding authors. Address: Barcelona Hepatic Hemodynamic Laboratory, Liver Unit, Hospital Clínic, Villarroel 170, Barcelona 08036, Catalonia, Spain. Tel.: +34 932-275-400 (2209), fax: +34 932-279-856 (V. Hernández-Gea), or Tel.: +34 932-275-400 (2209), fax number: +34 932-279-856 (J.C. García-Pagán).
E-mail addresses: vihernandez@clinic.cat (V. Hernández-Gea), jcgarcia@clinic.cat (J.C. García-Pagán).
† Both authors share first authorship.

https://doi.org/10.1016/j.jhep.2021.05.014

biological networks consider the relationship between changing genes as a whole. Weighted gene co-expression network analysis (WGCNA) is a multigene analysis that uses unsupervised hierarchical clustering (not depending on a priory defined gene set or pathways) to identify groups of genes commonly deregulated and group them into modules. WGCNA establishes gene co-expression modules, which are a cluster of genes that maintain a consistent expression pattern and possibly share a common biological regulatory role.[2] Therefore, WGCNA is an effective method to uncover the relationship between networks/genes and phenotypes.

In this unprecedented study, we used biological networks to better understand the biological pathology of PSVD by integrating transcriptomic and clinical data with the final aim of uncovering altered biological processes in patients with PSVD.

Patients and methods

Patients

The diagnosis of PSVD with portal hypertension was based on previously described criteria. Briefly[1,3]: i) presence of signs of portal hypertension (specific signs: gastroesophageal varices, porto-collateral circulation; non-specific signs: ascites, platelet count <150,000 per μl, splenomegaly); ii) absence of cirrhosis, advanced fibrosis or other causes of chronic liver diseases by appropriate serological and biochemical tests and by liver biopsy (performed in all patients of the cohort); and iii) a liver biopsy supporting the diagnosis of PSVD. Findings at liver biopsy supporting the diagnosis of PSVD are: obliterative portal venopathy (thickening of vessel wall, occlusion of the lumen, and vanishing of portal veins), nodular regenerative hyperplasia, incomplete septal fibrosis or cirrhosis, and non-specific signs like portal tract abnormalities (multiplication, dilation of arteries, periportal vascular channels, and aberrant vessels) or architectural disturbances (irregular distribution of the portal tracts and central veins, non-zonal sinusoidal dilation, mild perisinusoidal fibrosis), together with the absence of bridging fibrosis or cirrhosis and other significant pathologic abnormalities (inflammation, granulomas, steatosis, *etc.*).[1] Only patients with unequivocal diagnosis of PSVD[1] defined by either 1 sign specific for portal hypertension or 1 histological lesion specific for PSVD, or 1 sign not specific for portal hypertension and 1 histological lesion not specific for PSVD, were included in the study.[1]

Although, the current PSVD diagnostic criteria do not include the exclusion of portal vein thrombosis, thrombosis of both the hepatic and portal veins were discarded by imaging studies performed at diagnosis in all patients included in the current study.

Liver biopsies were performed by transjugular or percutaneous route and were revaluated by an experienced pathologist (AD) for the purpose of the study. Briefly, the specimens were formalin fixed, embedded in paraffin and cut into 4 μm sections. Hematoxylin-eosin, Masson's trichrome and reticulin staining were performed. Only specimens containing ≥6 complete portal tracts were considered suitable for diagnosis.[4] Those patients with liver biopsies including <6 portal tracts were not included in the study because the diagnosis of PSVD was not definite.

Patients with cirrhosis and portal hypertension and individuals with histologically normal livers (HNL) were included as control groups. Patients with cirrhosis and clinically significant portal hypertension, defined by hepatic venous pressure gradient ≥10 mmHg and/or presence of gastroesophageal varices, were retrospectively selected from the hepatic hemodynamic laboratory patient database and matched by sex, age and the severity of liver dysfunction with patients with PSVD. Adults with HNL and no history of receiving any treatment known to induce liver alterations (*i.e.* azathioprine, oxaliplatin, *etc.*) were retrospectively selected from the pathology database.

Twenty consecutive patients with PSVD – diagnosed from 2001 to 2010 and from whom remnants of the diagnostic liver biopsies were available – were included as an initial cohort. Twenty-one patients with cirrhosis and 13 HNL were used as controls. Another group of 8 patients with PSVD, 8 with cirrhosis and 8 with HNL were used as a validation cohort.

The protocol was reviewed and approved by the ethical committee at our institution (HCB/2009/5448). All patients registered in our database gave specific informed consent to use their clinical data and the remnants of their liver biopsies for research studies approved by the Ethical Committee.

Transcriptomic analysis in the initial cohort

RNA samples: we used formalin-fixed, paraffin embedded (FFPE) liver tissue from patients with PSVD, cirrhosis and HNL. RNA was extracted from 10 μm sections using the RNeasy FFPE KIT (QIAGEN GmgH, Hilden, Germany) according to manufacturer's instructions. Briefly, punch biopsies, 0.6 mm in diameter, were obtained from each tissue block and were placed in a tube containing 1 ml xylene. Immediately after the 100% ethanol extraction step, an additional 70% ethanol extraction was performed. Proteinase K digestions were at 55 °C for 15 min followed by 15 min at 80 °C. Genomic DNA was removed by DNaseI treatment and RNA purified using RNeasy MinElute Spin columns and eluted in a volume of 14–30 μl. For the validation cohort, RNA was extracted from frozen liver biopsy remnants with Trizol (Promega) according to manufacturer's extraction.

Microarray experiments were performed employing Whole-Genome DASL HT Assay covering 29,377 probes derived from the NCBI RefSeq database. RNA was extracted and analyzed from the following liver tissue samples: 20 PSVD, 21 cirrhosis and 13 HNL. RNA samples were processed, hybridized onto Beadchip arrays and analyzed using a BeadArray Reader (Illumina). Microarray expression data was computed using BeadStudio data analysis software (Illumina) and subsequently processed employing quantiles normalization using the Lumi bioconductor package. We selected those probes with a detection p value lower than 0.01 in at least 10% of samples, which resulted in a list of 24,578 evaluable probes, corresponding to 18,303 genes. For the detection of differentially expressed genes, a linear model was fitted to the data and empirical Bayes moderated statistics were calculated using the limma R package. p values were adjusted by determining false discovery rates using Benjamini-Hochberg procedure to identify genes differentially expressed between liver tissue samples from PSVD and the control group (those with cirrhosis and HNL).[5]

RNA sequencing studies

In the validation cohort we used RNA extracted from frozen liver biopsy remnants from 8 patients with PSVD, 8 with cirrhosis and 8 with HNL. Libraries were synthetized with the TruSeq stranded mRNA Library Prep (Illumina) and sequenced 2 * 125 bp on Illumina's HiSeq2500, obtaining a minimum of 35 million reads per sample. STAR program[6] against Homo Sapiens genome (hg38) was used for mapping the reads followed by the

quantification of genes with the RSEM[7] program using GENCODE v26 reference annotation.[8] Afterwards, to eliminate genes with at least an expected value greater than 5, we used TMM method and limma-voom transformation[9] to normalize the non-biological variability. Differential expression between different groups was assessed using moderated t-statistics.[5] Canonical pathway enrichment analysis was performed through GSEA function in clusterProfiler package using previously computed t-statistics.

All liver tissue microarray and RNA sequencing data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO; accession numbers GSE77627 and GSE171248, respectively).

Bioinformatics and statistical analysis

To explore the correlation patterns among genes across the data derived from the 57 protocol biopsies included in the initial cohort we employed the WGCNA R package as described by Langfelder and Horvath.[10] The PickSoftThreshold function was used to select the soft threshold power used to construct a network based on the criterion of approximate scale-free topology.

The power value selected was the lowest power for which the scale-free topology fit index curve flattened out upon reaching a value of 0.80. For network construction and module detection, the blockwise Module function was used with the following parameters: corType="pearson", networkType="signed", mergeCutHeight = 0.25, minModuleSize = 20). We computed the eigengene for each module, defined as the first principal component of the module representing the overall expression level of the module, as well as module membership (MM). We quantified associations of individual genes with our clinical variables by defining gene significance (GS) as (the absolute value of) the correlation between the gene and the trait. For each module, we also defined a quantitative measure of MM as the correlation of the module eigengene and the gene expression profile. This allowed us to quantify the similarity of all genes on the array to every module.

We assessed the association of a module to a PSVD phenotype as the number of genes statistically upregulated by means of a linear model comparing PSVD *vs.* the rest using limma R package. Values were transformed to GS parameter using the –log (moderated *p* value). To identify canonical pathways enriched in modules of co-expressed genes we employed the enricher function from ClusterProfiler R package.[11] The genesets employed to find enrichment in our clusters were obtained from MsigDB.[5]

Results

At the time of sampling, all patients included in the initial cohort with PSVD had clinical signs of portal hypertension. The most frequent sign of portal hypertension in patients with PSVD was splenomegaly followed by presence of gastroesophageal varices. Splenomegaly was present in 16 (80%) patients with a mean size of 15 ± 2.4 cm. Mean hepatic venous pressure gradient (HVPG) was 8 ± 4 mmHg. Patients with cirrhosis had similar liver function as those with PSVD (76% Child-Pugh A and 24% Child-Pugh B) with a mean HVPG of 12 ± 5 mmHg. Splenomegaly was present in 10 (47%) patients and median spleen size was 11 ± 24 cm (Table 1). Table S1 shows clinical and histological data for each of the patients with PSVD, drug exposure and associated diseases.

Identification of gene expression modules differentially expressed in PSVD

In the initial cohort, we used WGCNA to investigate correlation patterns among genes across microarray liver tissue samples from patients with PSVD, cirrhosis and HNL that might jointly contribute to PSVD development. WGCNA identified 12 clusters (modules), corresponding to a total of 3,637 highly correlated genes (from a total of 18,303 genes) in livers from patients with PSVD, cirrhosis and HNL (Fig. 1A). The number of genes for each module is summarized in Table 2, including the number of genes with fold change ≥1.5 or <-1.5 in PSVD compared with the combination of cirrhosis and HNL and with a *p* value <0.05. Information about the membership of these 3,637 genes is included in Table S2.

Because we were interested in the modules whose expression is coordinated in a similar manner in liver tissue samples from patients with PSVD, but which differs from expression in the control group (cirrhosis plus HNL), we compared the expression

Table 1. Baseline characteristics.

	PSVD (n = 20)	Cirrhosis (n = 21)	HNL group (n = 13)
Age (years)	42 ± 17	52 ± 9	50 ± 17
Sex (male)	13 (65%)	13 (62%)	6 (46%)
Presence GEV	13 (65%)	11 (52%)	n.a.
HVPG (mmHg)	7 ± 3	14 ± 4	n.a.
Splenomegaly	16 (80%)	10 (47%)	n.a.
Spleen size (cm)	15 ± 2.4	11 ± 2.4	n.a.
Child-Pugh score (A/B/C)	(15/5/0)	(16/5/0)	n.a.
Laboratory data			
Platelet count (x10^9/L)	150 ± 131	151 ± 185	273 ± 109
Bilirubin (mg/dl)	1.4 ± 1.2	1.5 ± 1.2	0.8 ± 0.4
Hemoglobin (g/L)	12.5 ± 2.3	13.0 ± 2.5	13.3 ± 1.9
Creatinine (mg/dl)	1.08 ± 0.8	0.9 ± 0.15	0.9 ± 0.13
AST (U/L)	47 ± 32	104 ± 69	29 ± 10
ALT (U/L)	41 ± 22	108 ± 74	36 ± 31
Albumin (g/L)	39.3 ± 4.6	37.1 ± 9	42 ± 3.9
PT (%)	78 ± 13	77 ± 10	95 ± 10
Liver stiffness (kPa)	10.5 ± 7.1	20.9 ± 9.9	n.a.

*Mean ±SD.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GEV, gastroesophageal varices; HNL, histologically normal liver; HVPG, hepatic venous pressure gradient; PSVD, porto-sinusoidal vascular disease; PT, prothrombin time.

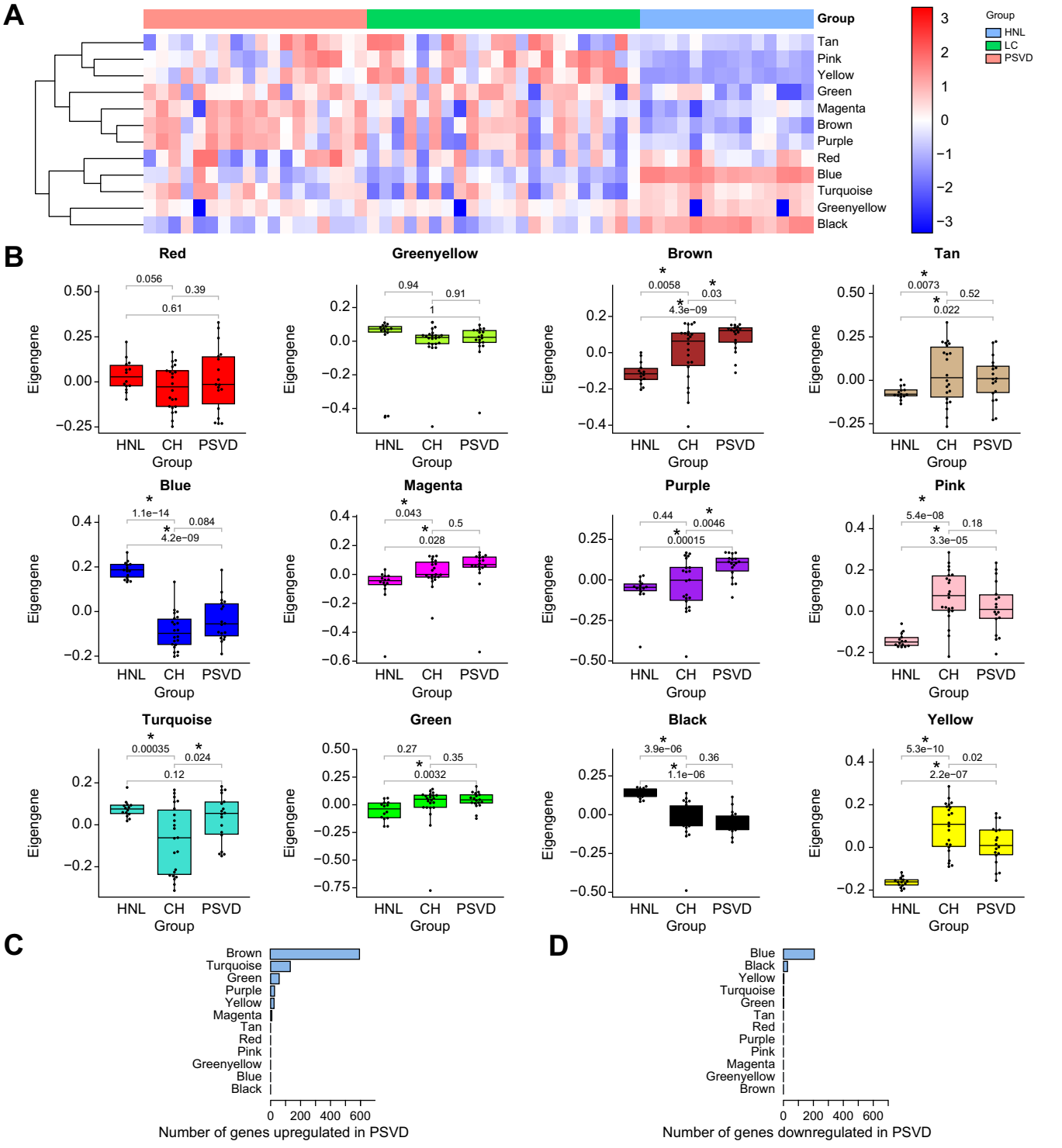


Fig. 1. Weighted gene co-expression network analysis of PSVD, cirrhosis and HNL samples. (A) Heatmap displaying the eigengene expression levels of each sample for each of the 12 modules identified by WGCNA. (B) Box and Whisker plots representing the Eigengene expression median in each module of PSVD, CH and HNL samples. Differences between groups were assessed by Student's *t* test. **p* value <0.05. (C-D) Bar plots showing the number of significantly upregulated and downregulated genes, respectively, in PSVD livers when compared to the combination of CH and HNL livers in each of the 12 modules detected by WGCNA. Differences between groups were assessed by Student's *t* test. A gene was considered up- or downregulated in PSVD when its fold-change *vs.* CH+HNL was >1.5 or <-1.5 and its *p* value <0.05. CH, cirrhosis; HNL, histologically normal liver; PSVD, porto-sinusoidal vascular disease; WGCNA, weighted gene correlation network analysis.

Table 2. Number of genes with significant fold change (overexpressed or underexpressed) in patients with PSVD *vs.* controls (cirrhosis plus HNL).

Module	Genes in the module with Fc ≥1.5 or <-1.5 & p <0.05	Total genes in the module
Pink	0	40
Brown	596	749
Yellow	26	610
Turquoise	132	898
Green	58	331
Blue	206	806
Magenta	8	38
Red	0	51
Green-yellow	0	23
Black	27	41
Purple	26	28
Tan	0	22

Significance is calculated by Student's t test.
Fc, fold change; HNL, histologically normal liver; PSVD, porto-sinusoidal vascular disease.

of eigengenes (genes representing the average expression of each module) from PSVD samples to the expression of cirrhosis and HNL in every module (Fig. 1B); only the Purple and Brown modules showed significance in both comparisons, both showing an upregulation of the eigengenes in the PSVD samples when compared to cirrhosis and HNL. However, since the Purple module consisted of only 28 genes, we focused on the Brown module, which consisted of 749 genes and showed the greater number of statistically differentially expressed genes, with 596 of the 749 genes being upregulated (fold change >1.5; p value <0.05) and none significantly downregulated (Fig. 1C-D).

We then explored the relationship of the transcriptomic modules with clinical variables that may have a differential profile in patients with PSVD compared to patients with cirrhosis and HNL, such as HVPG values, presence of esophageal varices, platelet count, bilirubin levels, spleen size and liver elastography values. The relationship between all the modules and the clinical features are shown in a correlation heatmap, in which the Y-axis corresponds to groups of genes (modules) and the x-axis includes the clinical variables of interest (Fig. 2A).

We next plotted the scatterplot of gene significance (GS) *vs.* MM in the co-expression Brown module. GS represents the

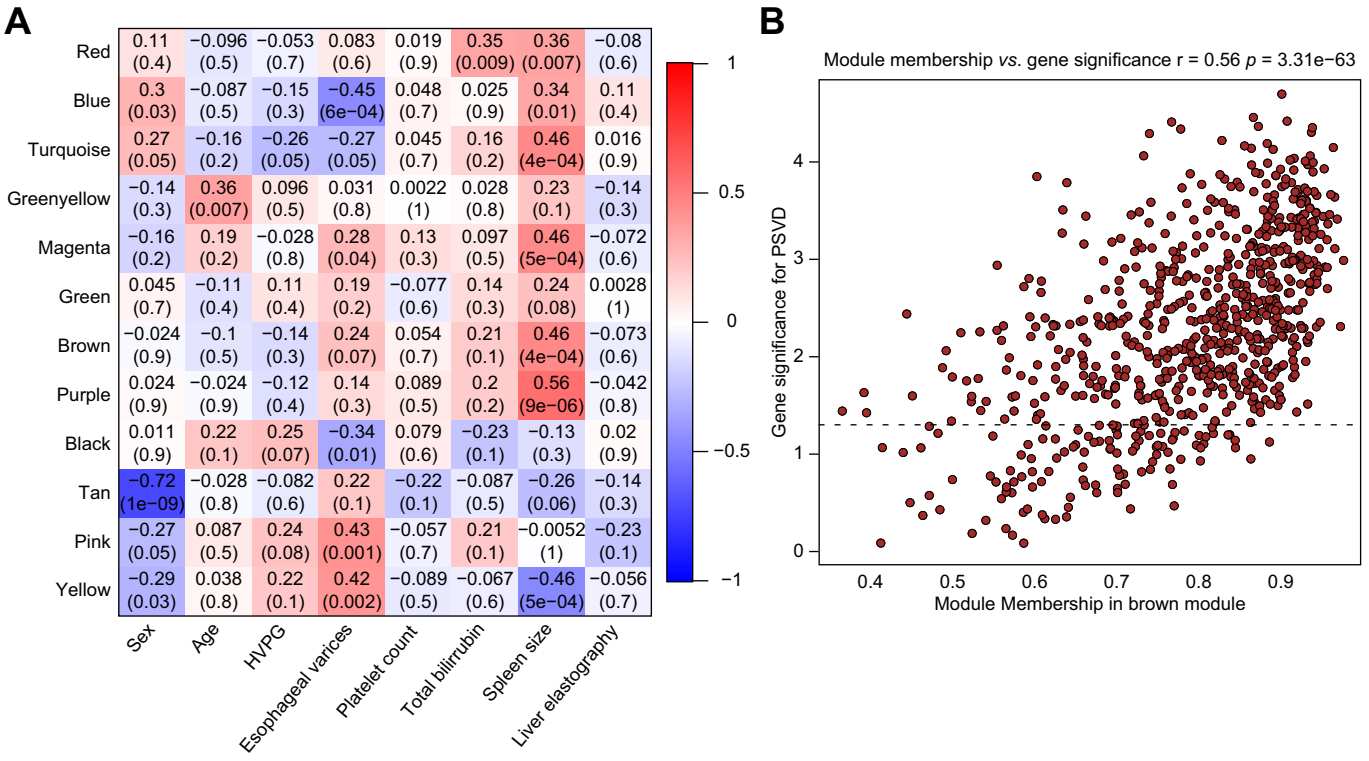


Fig. 2. Correlation of clinical variables with each of the WGCNA modules. (A) Heat map showing the Pearson correlation of clinical variables (x-axis) with each of the 12 modules (y-axis) identified by WGCNA. Associated p values were calculated by the cor.test R function. The color scale in the heat map corresponds to the magnitude of the Pearson correlation coefficients. Numbers in each field represent correlation and p value (in brackets). (B) Scatterplot of gene significance *vs.* module membership of genes of the Brown module. Genes from the brown module showed a significant correlation with gene significance (correlation 0.56; p <0.002). HVPG, hepatic venous pressure gradient; WGCNA, weighted gene correlation network analysis.

correlation between individual genes and the clinical traits associated with PSVD. MM is calculated by the WGCNA and quantifies how close a specific gene is to a given module. Fig. 2B shows that GS and MM are highly associated, implying that hub genes of the Brown module are significantly correlated with PSVD clinical characteristics and may be involved in PSVD physiopathology. Considering all these results, we chose the co-expression Brown module for further analysis.

Characterization of Brown module content and hub genes

In order to understand the physiological processes captured by each module, we firstly evaluated the molecular pathways significantly enriched in all the different modules (Table S3). Table 3 shows the 12 most enriched pathways in the Brown module. Importantly, the majority of them belong to canonical pathways involved mainly in hemostasis and coagulation, The fact that they are enriched in PSVD samples when compared to cirrhosis and HNL suggest that alteration in their regulation may be a pathogenic driving event occurring specifically or with higher intensity in PSVD.

To confirm the importance of the identified pathways in PSVD, we validated our results in a different cohort of patients (validation cohort). We analyzed 8 patients with PSVD diagnosed between years 2016-2019 (individual patient data in Table S1), together with 8 age-matched individuals with HNL and 8 patients with cirrhosis matched by age and liver function (baseline characteristics in Table S4), and performed RNA sequencing of stored frozen liver biopsy remnants from these patients (gene expression data in Table S5). We then performed gene set enrichment analysis in this new cohort comparing PSVD samples *vs.* the combination of both cirrhosis and HNL and analyzed the status of the 12 most significantly upregulated pathways from the Brown module (Table S6). As observed in Fig. 3, 9 of these 12 pathways were also statistically enriched in patients with PSVD from this new cohort, hence, validating our initial discoveries. This data highlights the involvement of hemostasis and coagulation but also lipid metabolism and oxidative phosphorylation in the development of PSVD.

We next analyzed the core-related genes (hub genes) from the 9 corroborated enriched pathways originally identified in the Brown module, and integrated the information on interactions between them by creating a network map. Graphically, genes are represented as nodes and their interactions are modeled as links between nodes. Color intensity is proportional to their corresponding level of connectivity (Fig. 4). Interestingly, 2 clusters of interconnected pathways were identified. The most enriched cluster integrated pathways related to 'Complement and coagulation cascades', 'Extrinsic pathway', 'Integrin pathway' and 'Formation of fibrin clot cascade', clearly suggesting that altered control of hemostasis and coagulation might be an essential trigger of, or contributor to, PSVD development. 'Oxidative phosphorylation', 'Tricarboxylic acid cycle and respiratory electron transport' and 'Mitochondrial protein import' grouped together in a smaller cluster, suggesting that mitochondrial oxidative phosphorylation may also play a role in PSVD.

Next, we focused on the hub genes linking at least 2 disrupted pathways and identified the Serpin family (*SERPINC1*), the apolipoproteins (*APOA*, *APOB*, *APOC*), ATP synthases (*ATP5G1*, *ATP5B*), fibrinogen genes (*FGB*, *FGA*) and alpha-2-macroglobulin (*A2M*) as highly connective genes that may have an important role in PSVD pathogenesis.

SERPIN genes are potent protease inhibitors able to modulate both thrombotic and thrombolytic pathway activation, which are essential for blood clotting. These fibrinolytic regulators are considered to play an important role in the maintenance of vascular homeostasis and regulation of vascular remodeling. Deregulation in serpin family genes has been linked to cardiovascular disease, endothelial dysfunction and vascular remodeling.[12–15]

Apolipoproteins are mostly synthesized in the liver and are components of chylomicrons, VLDL, IDL, and LDL. Similar to Serpins, increased levels have been linked to endothelial damage, cellular inflammation,[16] platelet activation and increased atherosclerosis risk.[17,18]

A2M encodes for a protease inhibitor protein essential in maintaining hemostatic balance. A2M deregulation has been associated with endothelial dysfunction, arteriosclerosis and cardiovascular risk.[19–21]

Genes encoding fibrinogen (*FGA* & *FGB*) play a critical role following vascular injury and their dysfunction has been strongly associated with cardiovascular risk, coronary heart disease and cardiovascular events.[22,23]

ATP5 genes encode the ATP synthase enzyme located mainly in the mitochondria where it catalyzes ATP during oxidative phosphorylation. However, ATP synthase can also be located in endothelial cells and regulates endothelial cell function and proliferation. Altered expression of ATP synthase provokes both

Table 3. Twelve top significant molecular pathways enriched in the Brown module.

Geneset	DB*	N**	Adjusted *p* value
Complement and coagulation cascades	KEGG	16	1,46E-08
Response to elevated platelet cytosolic Ca^{2+}	Reactome	15	3,29E-06
Formation of fibrin clot clotting cascade	Reactome	8	3,61E-05
Oxidative phosphorylation	KEGG	17	4,11E-05
Extrinsic Prothrombin Activation Pathway	Biocarta	5	1,18E-04
Mitochondrial protein import	Reactome	10	1,19E-04
HDL-mediated lipid transport	Reactome	5	2,56E-04
Intrinsic Prothrombin Activation Pathway	Biocarta	6	2,73E-04
TCA cycle and respiratory electron trasnport	Reactome	15	7,39E-04
Lipoprotein metabolism	Reactome	6	8,56E-04
Hemostasis	Reactome	34	9,28E-04
Beta2 integrin cell surface interactions	PID	6	1,04E-03

TCA, tricarboxylic acid.
*DB = Pathway database.
**N = Number of genes present in the module.

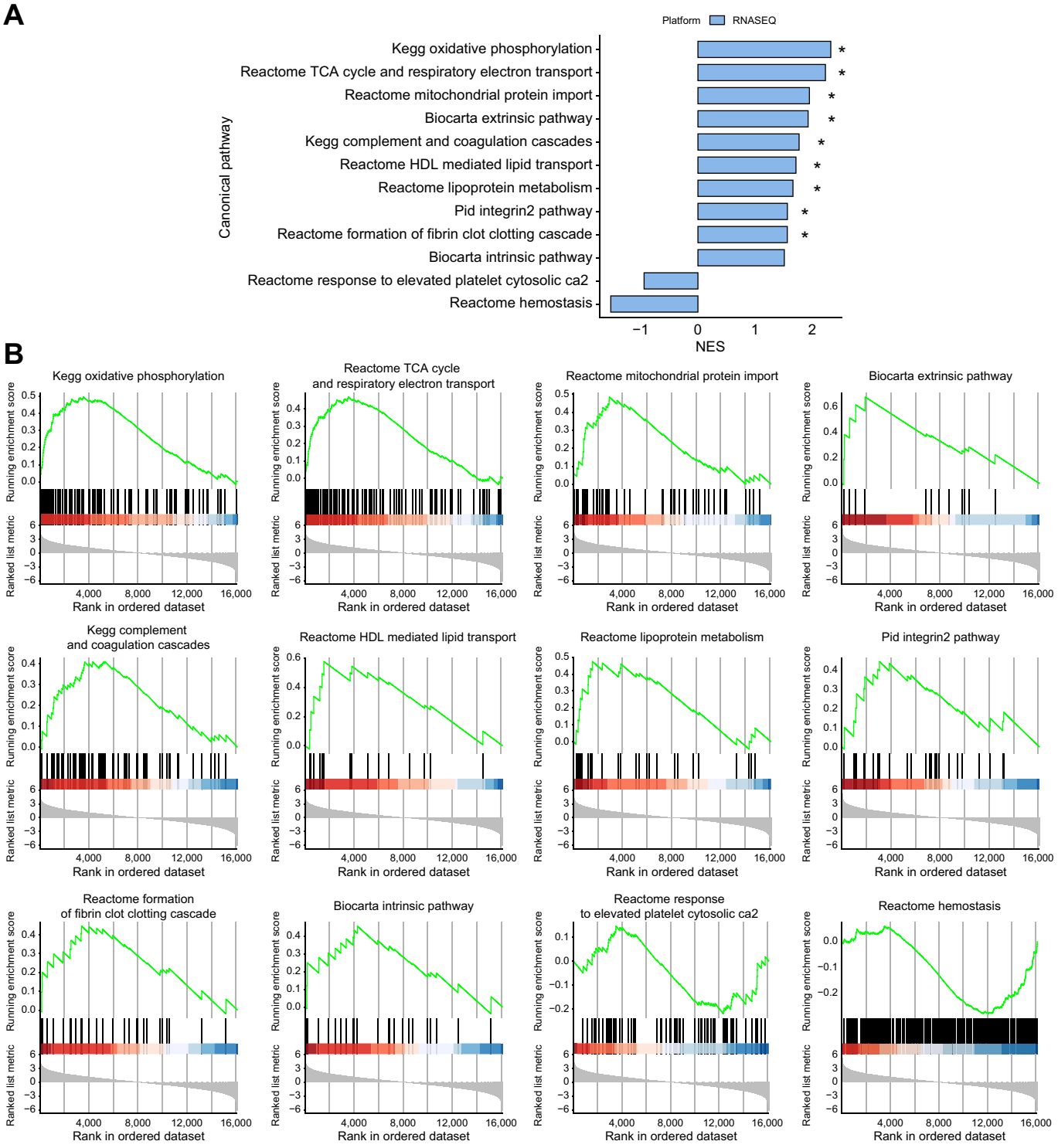


Fig. 3. Validation of selected pathways in a new cohort of patients by GSEA. GSEA of the new cohort of 8 PSVD, 8 cirrhosis and 8 HNL samples. (A) Bar plots represent the NES for each of the 12 indicated pathways in patients with PSVD compared to the combination of cirrhosis and HNL. FDR (q value) is provided by GSEA function from clusterprofiler. *Indicates a *p* value <0.05 and an FDR <0.1. (B) GSEA from PSVD livers of the validation cohort when compared to controls (cirrhosis and HNL). Plots represent each of the 12 pathways originally identified as highly enriched in the Brown module. FDR, false discovery rate; GSEA, gene set enrichment analysis; HNL, histologically normal liver; NES, normalized enrichment score; PSVD, porto-sinusoidal vascular disease; TCA, tricarboxylic acid.

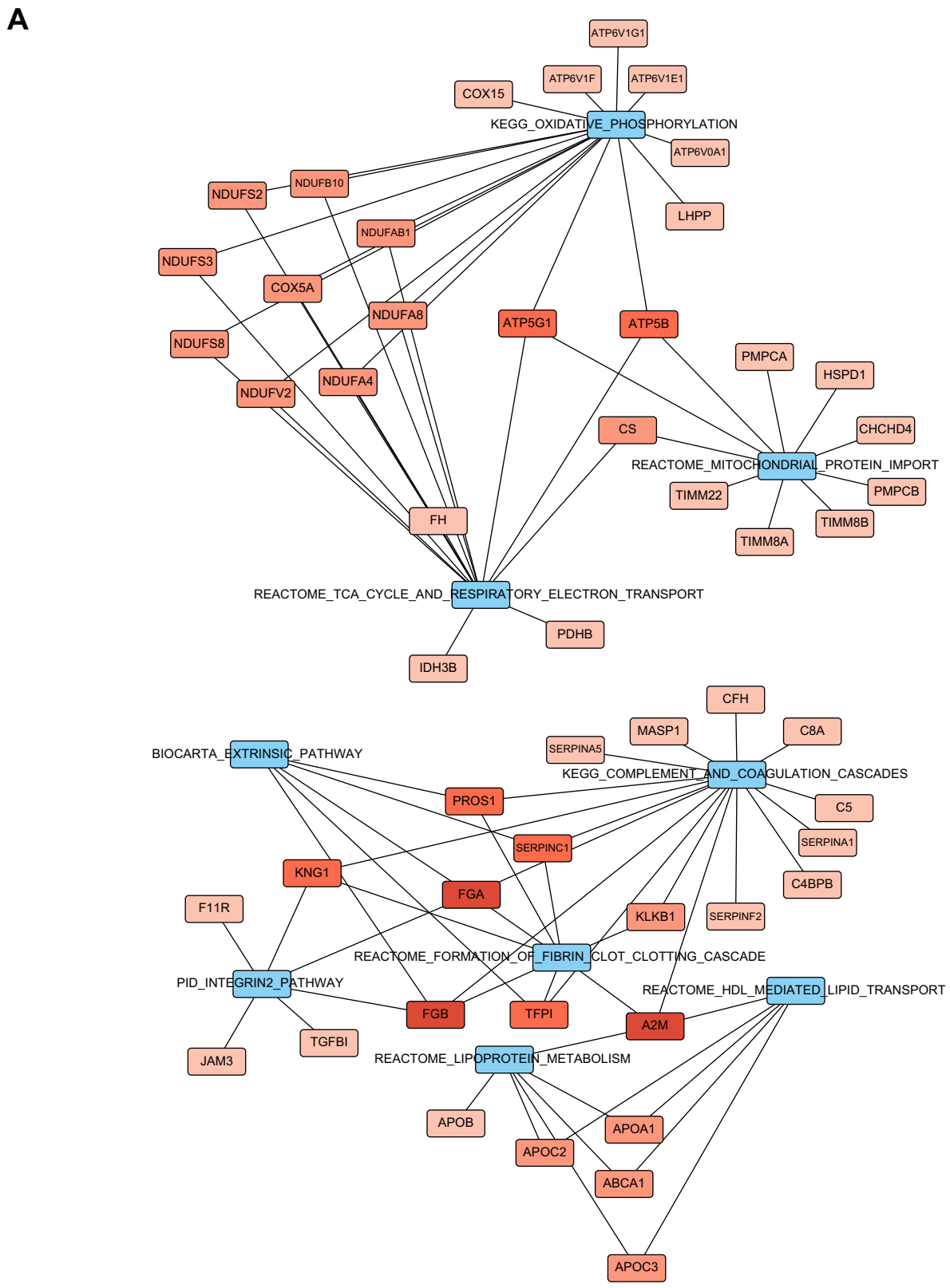


Fig. 4. Network enrichment analysis representation (Cnetplot) of genes in the Brown module. The connectome displays the 9 validated gene sets enriched in PSVD patients *vs.* HNL and cirrhosis. Pathways involved are represented in blue and individual genes in red. Color intensity is proportional to their corresponding level of connectivity with different pathways. HNL, histologically normal liver; PSVD, porto-sinusoidal vascular disease.

mitochondrial dysfunction and vascular inflammation and alteration of the vascular tone leading to cardiovascular disease.[24–26]

In conclusion, these results suggest that patients with PSVD have a genetic profile distinct from those with cirrhosis. This profile is mostly enriched in genes involved in maintaining vascular homeostasis, pointing to a potential central role of the liver endothelium in the pathogenesis of PSVD.

Discussion

The pathogenesis of PSVD is unknown and we lack curative treatments able to modify the natural history and prevent the development of portal hypertension. Disorders of immunity, prothrombotic conditions, infections, drug exposure and familiar aggregation have been documented in patients with PSVD, although a causal relation has only been established in the latter.[27–29] Studies aimed at further clarifying the underlying disease mechanisms are needed in order to identify new therapeutic targets.

The aim of our study was to identify the signaling pathways specifically deregulated in patients with PSVD compared with cirrhosis and HNL. To the best of our knowledge, this is the first study using a systems biology strategy to identify, in an unsupervised manner, the transcriptomic profile of patients with PSVD. We decided not to use traditional differential expression analysis based on predefined gene sets of pathways but a modular analysis that explored the correlation between groups of co-expressed genes using an unsupervised hierarchical clustering approach. WGCNA analyzes the expression pattern of multiple sample genes adopting a weighted expression strategy and not a scale distribution. WGCNA allows for the identification of interactions between genes and can reflect how well a given gene is connected to other genes with the main advantage of identifying pathways jointly involved in a biological process. For this purpose, WGCNA divides gene co-expression networks of complex biological processes into modules. Analysis of the modules together with the clinical traits of the disease can identify key regulatory pathways involved in the pathogenesis of the disease.[30] We identified 12 different modules by WGCNA, from which the Brown module exhibited the greatest positive correlation with PSVD and contained the highest number of differentially deregulated genes. This initial analysis reinforced the hypothesis that PSVD has a unique genetic signature, supporting the pre-existing idea that the mechanisms implicated in the disease development are different from those implicated in cirrhosis.

Enrichment analysis identified that genes in the Brown module were primarily associated with pathways related with vascular homeostasis. We then analyzed gene interactions and quantified network relationships between the altered biological processes in patients with PSVD. The top hub genes overexpressed in PSVD samples and linking altered pathways belonged to the Serpin family (*SERPINC1*), the apolipoproteins (*APOA*, *APOB*, *APOC3*), ATP synthases (*ATP5G1*, *ATP5B*), fibrinogen genes (*FGB*, *FGA*) and *A2M*.

Although not previously linked to PSVD, all have been previously related to vascular injury and endothelial cell dysfunction. Available evidence coming from different fields correlates deregulation in these genes individually with vascular processes such as vascular remodeling, atherosclerosis and endothelial dysfunction. Of interest, vascular damage is a common finding in the liver of patients with PSVD. Indeed, obliteration of the small and medium intrahepatic vascular branches is a major diagnostic histologic criterion. Deregulation of pathways involved in vascular homeostasis may alter vascular remodeling of the intrahepatic vascular branches causing the common histological findings of PSVD (obliterative portal venopathy, sclerosis of small and medium branches of the portal vein, venopenia, narrowing of vessel lumen). Whether these changes are initial drivers of the disease remains an open question as we decided to include only patients with an unequivocal diagnosis of PSVD and portal hypertension, *i.e.* those with advanced disease. Why these pathways are de-regulated in PSVD and whether there is a common upstream regulator to all of them cannot be answered with this exploratory study and deserves further research.

We only included patients with a certain diagnosis of PSVD (biopsy proven) and portal hypertension. Although the relatively low number of patients with PSVD enrolled could be considered a limitation of the study, since PSVD is a rare condition, a sample of 20 patients could be considered adequate. Indeed, with this number of patients, we were able to find a module of genes differentiating patients with PSVD. In addition, we were able to validate 9 of the 12 most enriched pathways of the Brown module in a new cohort of patients despite important differences between the 2 cohorts in terms of the quality of the starting material (FFPE *vs.* frozen tissue), the platform chosen for the analysis (microarray *vs.* RNA sequencing) and the reduced number of samples of the validation cohort, with only 8 samples per group and lower statistical power. Therefore, although this study was initially conceived as a proof-of-concept, the addition of this small validation cohort conferred more robustness to our findings and is indicative of the significance of the study.

In conclusion, our results reveal that PSVD has a unique transcriptomic profile and that deregulation of pathways involved in vascular homeostasis and oxidative phosphorylation are drivers of the disease. This study highlights the need to better understand liver vascular biology and mechanisms of vascular remodeling as a means to better understand the pathogenesis of PSVD. Meanwhile, although molecular biology studies aimed at improving our understanding of intrahepatic vascular damage are needed, maintaining a healthy liver endothelium could be a promising therapeutic strategy.

Abbreviations

A2M, alpha-2-macroglobulin; FFPE, formalin-fixed, paraffin embedded; GS, gene significance; HNL, histologically normal livers; HVPG, hepatic venous pressure gradient; MM, module membership; PSVD, porto-sinusoidal vascular disease; WGCNA, weighted gene correlation network analysis.

Financial support

This study was supported by the Instituto de Salud Carlos III FIS PI17/00398, the Ministry of Education and Science, Spain (SAF-2016-75767-R); Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR-SGR2017-517) a grant from Generalitat de Catalunya, Fondo Europeo de Desarrollo Regional (FEDER) and Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), funded by Instituto de Salud Carlos III. Marta Magaz is a recipient of a Río Hortega grant from

Instituto de Salud Carlos III. Pol Olivas has been funded by Contractes Clínic de Recerca "Emili Letang-Josep Font" 2020, granted by Hospital Clínic de Barcelona.

Conflict of interest

The authors involved in this study declare they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Conceptualization: VHG, JCGP, GC. Data acquisition: SS, FB, VPC, AB, PO, AD Statistical analysis: JJL, JS. Funding Acquisition: JCGP, VHG. Methodology: VHG, JCGP, JJL. Analysis and interpretation of data: VHG, GC, JCGP, JJL, JS. Writing – original draft: VHG, JCGP, JL; Writing – review & editing: VHG, FB, VPC, SS, AD, GC, RGD, PO, MM, AB, FT, MRG, JS, JJL, JCGP. Study supervision: VHG, JCGP, JJL. All authors reviewed and approved the final manuscript.

Data availability statement

All liver tissue microarray and RNAseq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO; accession numbers GSE77627 and GSE171248, respectively).

Acknowledgements

The authors thank Ángeles Falgà, Maria José Serrano, Pamela Vizcarra and Nini Bustos for expert nursing support, and Oliva Ros for the administrative assistance.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.05.014.

References

Author names in bold designate shared co-first authorship

[1] De Gottardi A, Rautou PE, Schouten J, Rubbia-Brandt L, Leebeek F, Trebicka J, et al. Porto-sinusoidal vascular disease: proposal and description of a novel entity. Lancet Gastroenterol Hepatol 2019;4:399–411. https://doi.org/10.1016/S2468-1253(19)30047-0.

[2] **Stuart JM**, **Segal E**, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. Science (80-) 2003;302:249–255. https://doi.org/10.1126/science.1087447.

[3] Hernández-Gea V, Baiges A, Turon F, Garcia-Pagán JC. Idiopathic portal hypertension. Hepatology 2018;68:2413–2423. https://doi.org/10.1002/hep.30132.

[4] Kalambokis G, Manousou P, Vibhakorn S, Marelli L, Cholongitas E, Senzolo M, et al. Transjugular liver biopsy - indications, adequacy, quality of specimens, and complications - a systematic review. J Hepatol 2007;47:284–294. https://doi.org/10.1016/j.jhep.2007.05.001.

[5] Ritchie ME, Phipson B, Wu D, 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. https://doi.org/10.1093/nar/gkv007.

[6] Dobin A, Gingeras TR. Mapping RNA-seq reads with STAR. Curr Protoc Bioinform 2015;51:11.14.1–11.14.19. https://doi.org/10.1002/0471250953.bi1114s51.

[7] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. Bioinform Impact Accurate Quantif Proteom Genet Anal Res 2014:41–74. https://doi.org/10.1201/b16589.

[8] Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 2019;47:D766–D773. https://doi.org/10.1093/nar/gky955.

[9] Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014;15:1–17. https://doi.org/10.1186/gb-2014-15-2-r29.

[10] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9. https://doi.org/10.1186/1471-2105-9-559.

[11] Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. Omi A J Integr Biol 2012;16:284–287. https://doi.org/10.1089/omi.2011.0118.

[12] Rietveld IM, Schreuder M, Reitsma PH, Bos MHA. Elevated coagulation factor levels affect the tissue factor-threshold in thrombin generation. Thromb Res 2018;172:104–109. https://doi.org/10.1016/j.thromres.2018.10.020.

[13] Slack MA, Gordon SM. Protease activity in vascular disease. Arterioscler Thromb Vasc Biol 2019;39:E210–E218. https://doi.org/10.1161/ATVBAHA.119.312413.

[14] Kuiper J, Quax P, Bot I. Anti-apoptotic serpins as therapeutics in cardiovascular diseases. Cardiovasc Hematol Disord Target 2015;13:111–122. https://doi.org/10.2174/1871529x11313020004.

[15] Richardson J, Viswanathan K, Lucas A. Serpins, the vasculature, and viral therapeutics. Front Biosci 2006;11:1042–1056. https://doi.org/10.2741/1862.

[16] **Tao Y**, **Xiong Y**, Wang H, Chu S, Zhong R, Wang J, et al. APOC3 induces endothelial dysfunction through TNF-α and JAM-1. Lipids Health Dis 2016;15:1–8. https://doi.org/10.1186/s12944-016-0326-0.

[17] Riwanto M, Rohrer L, Roschitzki B, Besler C, Mocharla P, Mueller M, et al. Altered activation of endothelial anti-and proapoptotic pathways by high-density lipoprotein from patients with coronary artery disease: role of high-density lipoprotein-proteome remodeling. Circulation 2013;127:891–904. https://doi.org/10.1161/CIRCULATIONAHA.112.108753.

[18] Khetarpal SA, Qamar A, Millar JS, Rader DJ. Targeting ApoC-III to reduce coronary disease risk. Curr Atheroscler Rep 2016;18. https://doi.org/10.1007/s11883-016-0609-y.

[19] Celermajer DS, Sorensen KE, Gooch VM, Miller OI, Sullivan ID, Lloyd JK, et al. Non-invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. Lancet 1992;340:1111–1115. https://doi.org/10.1016/0140-6736(92)93147-F.

[20] Nezu T, Hosomi N, Aoki S, Deguchi K, Masugata H, Ichihara N, et al. Alpha2-macroglobulin as a promising biomarker for cerebral small vessel disease in acute ischemic stroke patients. J Neurol 2013;260:2642–2649. https://doi.org/10.1007/s00415-013-7040-x.

[21] Lygirou V, Latosinska A, Makridakis M, Mullen W, Delles C, Schanstra JP, et al. Plasma proteomic analysis reveals altered protein abundances in cardiovascular disease. J Transl Med 2018;16:1–12. https://doi.org/10.1186/s12967-018-1476-9.

[22] Iacoviello L, Vischetti M, Zito F, Donati MB. Genes encoding fibrinogen and cardiovascular risk. Hypertension 2001;38:1199–1203. https://doi.org/10.1161/hy1101.099478.

[23] Reiner AP, Carty CL, Carlson CS, Wan JY, Rieder MJ, Smith JD, et al. Association between patterns of nucleotide variation across the three fibrinogen genes and plasma fibrinogen levels: the Coronary Artery Risk Development in Young Adults (CARDIA) study. J Thromb Haemost 2006;4:1279–1287. https://doi.org/10.1111/j.1538-7836.2006.01907.x.

[24] Fu Y, Zhu Y. Ectopic ATP synthase in endothelial cells: a novel cardiovascular therapeutic target. Curr Pharm Des 2011;16:4074–4079. https://doi.org/10.2174/138161210794519219.

[25] Yamamoto K, Shimizu N, Obi S, Kumagaya S, Taketani Y, Kamiya A, et al. Involvement of cell surface ATP synthase in flow-induced ATP release by vascular endothelial cells. Am J Physiol - Hear Circ Physiol 2007;293. https://doi.org/10.1152/ajpheart.01385.2006.

[26] **Viiri LE**, **Full LE**, Navin TJ, Begum S, Didangelos A, Astola N, et al. Smooth muscle cells in human atherosclerosis: proteomic profiling reveals differences in expression of Annexin A1 and mitochondrial proteins in carotid disease. J Mol Cell Cardiol 2013;54:65–72. https://doi.org/10.1016/j.yjmcc.2012.11.002.

[27] Koot BGP, Alders M, Verheij J, Beuers U, Cobben JM. A de novo mutation in KCNN3 associated with autosomal dominant idiopathic non-cirrhotic portal hypertension. J Hepatol 2016;64:974–977. https://doi.org/10.1016/j.jhep.2015.11.027.

[28] Vilarinho S, Sari S, Yilmaz G, Stiegler AL, Boggon TJ, Jain D, et al. Recurrent recessive mutation in deoxyguanosine kinase causes idiopathic non-cirrhotic portal hypertension. Hepatology 2016;63:1977–1986. https://doi.org/10.1002/hep.28499.
[29] Besmond C, Valla D, Hubert L, Poirier K, Grosse B, Guettier C, et al. Mutations in the novel gene FOPV are associated with familial autosomal dominant and non-familial obliterative portal venopathy. Liver Int 2018;38:358–364. https://doi.org/10.1111/liv.13547.
[30] Londoño MC, Souza LN, Lozano JJ, Miquel R, Abraldes JG, Llovet LP, et al. Molecular profiling of subclinical inflammatory lesions in long-term surviving adult liver transplant recipients. J Hepatol 2018;69:626–634. https://doi.org/10.1016/j.jhep.2018.04.012.