

# Increased serum miR-193a-5p during non-alcoholic fatty liver disease progression: Diagnostic and mechanistic relevance

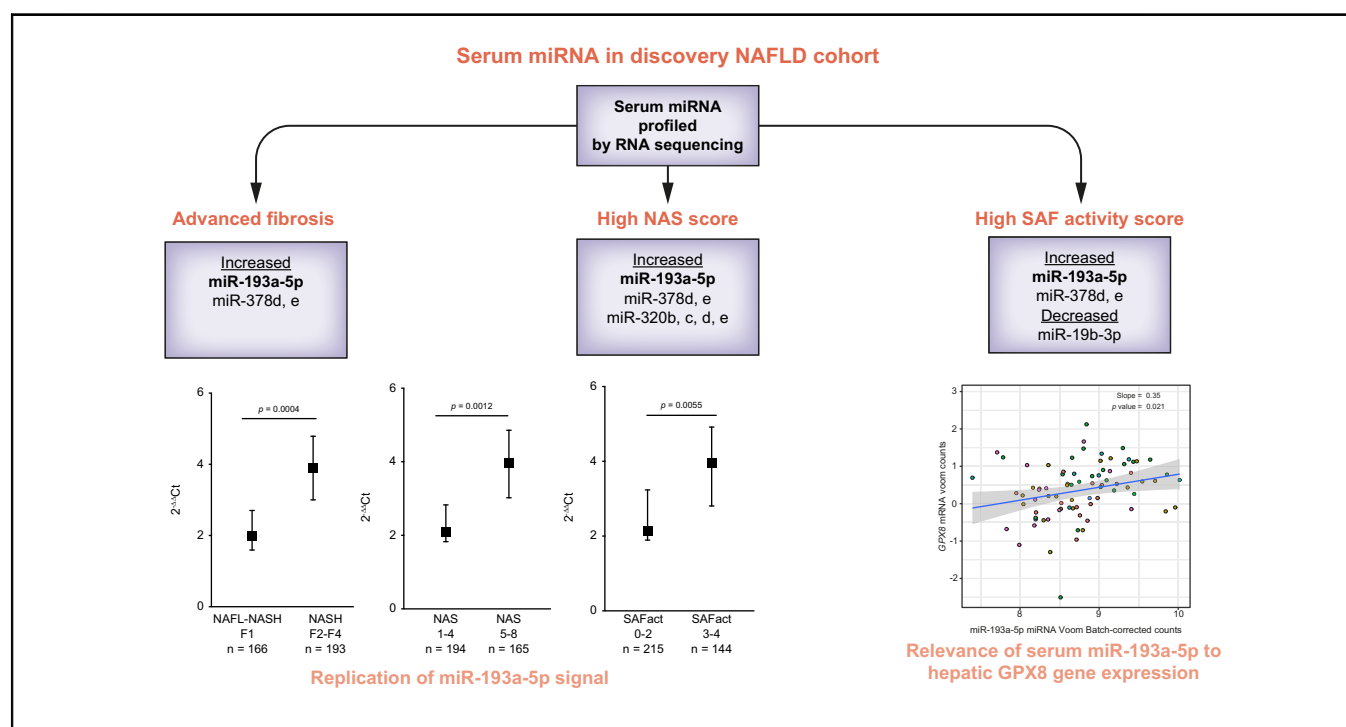
## Authors

Katherine Johnson, Peter J. Leary, Olivier Govaere, Matthew J. Barter, Sarah H. Charlton, Simon J. Cockell, Dina Tiniakos, Michalina Zatorska, Pierre Bedossa, M. Julia Brosnan, Jeremy F. Cobbold, Mattias Ekstedt, Guruprasad P. Aithal, Karine Clément, Jörn M. Schattenberg, Jerome Boursier, Vlad Ratziu, Elisabetta Bugianesi, Quentin M. Anstee, Ann K. Daly, on behalf of the LITMUS Consortium Investigators

## Correspondence

[a.k.daly@ncl.ac.uk](mailto:a.k.daly@ncl.ac.uk) (A.K. Daly), [quentin.anstee@newcastle.ac.uk](mailto:quentin.anstee@newcastle.ac.uk) (Q.M. Anstee).

## Graphical abstract



## Highlights

- Serum miRNA was sequenced in 183 NAFLD cases of varying severity and 10 population controls.
- Plasma levels of miR-193a-5p were significantly increased in patients with advanced fibrosis, high NAS scores, or high SAF scores.
- Other miRNAs including miR378d and miR378e were also significantly increased in certain comparisons.
- The findings for miR-193a-5p were replicated in a cohort of 372 additional NAFLD cases.

## Lay summary

MicroRNAs (miRNAs) are small pieces of nucleic acid that may turn expression of genes on or off. These molecules can be detected in the blood circulation, and their levels in blood may change in liver disease including non-alcoholic fatty liver disease (NAFLD). To see if we could detect specific miRNA associated with advanced stages of NAFLD, we carried out miRNA sequencing in a group of 183 patients with NAFLD of varying severity together with 10 population controls. We found that a number of miRNAs showed changes, mainly increases, in serum levels but that 1 particular miRNA miR-193a-5p consistently increased. We confirmed this increase in a second group of cases with NAFLD. Measuring this miRNA in a blood sample may be a useful way to determine whether a patient has advanced NAFLD without an invasive liver biopsy.



# Increased serum miR-193a-5p during non-alcoholic fatty liver disease progression: Diagnostic and mechanistic relevance

Katherine Johnson,<sup>1</sup> Peter J. Leary,<sup>2</sup> Olivier Govaere,<sup>1</sup> Matthew J. Barter,<sup>3</sup> Sarah H. Charlton,<sup>3</sup> Simon J. Cockell,<sup>2,3</sup> Dina Tiniakos,<sup>1</sup> Michalina Zatorska,<sup>1</sup> Pierre Bedossa,<sup>1</sup> M. Julia Brosnan,<sup>4,†</sup> Jeremy F. Cobbold,<sup>5</sup> Mattias Ekstedt,<sup>6</sup> Guruprasad P. Aithal,<sup>7</sup> Karine Clément,<sup>8,9</sup> Jörn M. Schattenberg,<sup>10</sup> Jerome Boursier,<sup>11</sup> Vlad Ratziu,<sup>8,9</sup> Elisabetta Bugianesi,<sup>12</sup> Quentin M. Anstee,<sup>1,13,\*‡</sup> Ann K. Daly<sup>1,\*‡</sup>, on behalf of the LITMUS Consortium Investigators<sup>§</sup>

<sup>1</sup>Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; <sup>2</sup>Bioinformatics Support Unit, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; <sup>3</sup>Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; <sup>4</sup>Internal Medicine Research Unit, Pfizer Inc., Cambridge, MA, USA; <sup>5</sup>Oxford Liver Unit, NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK; <sup>6</sup>Division of Diagnostics and Specialist Medicine, Department of Health, Medicine, and Caring Sciences, Linköping University, Linköping, Sweden; <sup>7</sup>NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, Nottingham, UK; <sup>8</sup>Institute of Cardiometabolism and Nutrition, Pitié Salpêtrière Hospital, Paris, France; <sup>9</sup>Assistance Publique – Hôpitaux de Paris, Paris, France; <sup>10</sup>Metabolic Liver Research Program, I. Department of Medicine, University Medical Center of Johannes Gutenberg-University Mainz, Mainz, Germany; <sup>11</sup>Hepatology Department, Angers University Hospital, Angers, France; <sup>12</sup>Division of Gastroenterology, Department of Medical Sciences, University of Turin, Turin, Italy; <sup>13</sup>Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Trust, Newcastle upon Tyne, UK

JHEP Reports 2022. <https://doi.org/10.1016/j.jhepr.2021.100409>

**Background & Aims:** Serum microRNA (miRNA) levels are known to change in non-alcoholic fatty liver disease (NAFLD) and may serve as useful biomarkers. This study aimed to profile miRNAs comprehensively at all NAFLD stages.

**Methods:** We profiled 2,083 serum miRNAs in a discovery cohort (183 cases with NAFLD representing the complete NAFLD spectrum and 10 population controls). miRNA libraries generated by HTG EdgeSeq were sequenced by Illumina NextSeq. Selected serum miRNAs were profiled in 372 additional cases with NAFLD and 15 population controls by quantitative reverse transcriptase PCR.

**Results:** Levels of 275 miRNAs differed between cases and population controls. Fewer differences were seen within individual NAFLD stages, but miR-193a-5p consistently showed increased levels in all comparisons. Relative to NAFL/non-alcoholic steatohepatitis (NASH) with mild fibrosis (stage 0/1), 3 miRNAs (miR-193a-5p, miR-378d, and miR378d) were increased in cases with NASH and clinically significant fibrosis (stages 2–4), 7 (miR193a-5p, miR-378d, miR-378e, miR-320b, miR-320c, miR-320d, and miR-320e) increased in cases with NAFLD activity score (NAS) 5–8 compared with lower NAS, and 3 (miR-193a-5p, miR-378d, and miR-378e) increased but 1 (miR-19b-3p) decreased in steatosis, activity, and fibrosis (SAF) activity score 2–4 compared with lower SAF activity. The significant findings for miR-193a-5p were replicated in the additional cohort with NAFLD. Studies in Hep G2 cells showed that following palmitic acid treatment, miR-193a-5p expression decreased significantly. Gene targets for miR-193a-5p were investigated in liver RNAseq data for a case subgroup (n = 80); liver GPX8 levels correlated positively with serum miR-193a-5p.

**Conclusions:** Serum miR-193a-5p levels correlate strongly with NAFLD activity grade and fibrosis stage. MiR-193a-5p may have a role in the hepatic response to oxidative stress and is a potential clinically tractable circulating biomarker for progressive NAFLD.

**Lay summary:** MicroRNAs (miRNAs) are small pieces of nucleic acid that may turn expression of genes on or off. These molecules can be detected in the blood circulation, and their levels in blood may change in liver disease including non-alcoholic fatty liver disease (NAFLD). To see if we could detect specific miRNA associated with advanced stages of NAFLD, we carried out miRNA sequencing in a group of 183 patients with NAFLD of varying severity together with 10 population controls. We found that a number of miRNAs showed changes, mainly increases, in serum levels but that 1 particular miRNA miR-193a-5p consistently increased. We confirmed this increase in a second group of cases with NAFLD. Measuring this miRNA in a blood sample may be a useful way to determine whether a patient has advanced NAFLD without an invasive liver biopsy.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: MicroRNA; Non-alcoholic fatty liver disease; Biomarker; Sequencing.

Received 10 June 2021; received in revised form 8 October 2021; accepted 9 November 2021; available online 25 November 2021

† Deceased.

‡ These authors contributed equally.

\* Corresponding authors. Address: Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Fourth Floor, William Leech Building, Framlington Place, Newcastle upon Tyne NE2 4HH, UK, Tel.: +44-191-208-7031 (A.K. Daly); Tel.: +44-191-208-7012 (Q.M. Anstee). E-mail addresses: [a.k.daly@ncl.ac.uk](mailto:a.k.daly@ncl.ac.uk) (A.K. Daly), [quentin.anstee@newcastle.ac.uk](mailto:quentin.anstee@newcastle.ac.uk) (Q.M. Anstee).



## Introduction

Non-alcoholic fatty liver disease (NAFLD) affects approximately one-quarter of the adult general population and is the most common cause of liver disease in the Western world.<sup>1</sup> The complex, multisystem disease is characterised by intra-hepatic triglyceride accumulation in the absence of excess alcohol intake.<sup>1,2</sup> In a significant minority of patients, NAFLD can progress from steatosis (non-alcoholic fatty liver [NAFL]) to its inflammatory form (non-alcoholic steatohepatitis [NASH]) with hepatocyte injury and progressive fibrosis and may lead to cirrhosis, hepatocellular carcinoma, and end-stage liver disease.<sup>3</sup>

NAFLD pathogenesis is multifactorial, and so, in addition to metabolic comorbidities and environmental influences, genetic and epigenetic factors confer an increased risk of NAFLD-associated end-stage liver disease.<sup>4,5</sup> Inflammatory disease severity may be assessed histologically using 1 of 2 widely adopted semiquantitative scoring systems: the NASH Clinical Research Network NAFLD activity score (NAS), which conflates the degree of steatosis with steatohepatitic activity grade, and the fatty liver inhibition of progression (FLIP) steatosis–activity–fibrosis (SAF) score, which preserves the distinction between steatosis and activity.<sup>6,7</sup> Histological scoring of a liver biopsy remains the current reference standard for grading and staging NAFLD for clinical trials, even though sampling error as well as inter- and intra-observer discrepancies are well recognised to increase variability and result in misclassification.<sup>8,9</sup> Consequently, circulating biomarkers have been sought to circumvent the need for invasive biopsies.

MicroRNAs (miRNAs) are small (~22 nucleotides in length) non-coding RNA molecules that can post-transcriptionally regulate gene expression.<sup>10</sup> Mature miRNAs, complexed with facilitative proteins (*i.e.* the Argonaute protein family), exert their effects by binding mRNA molecules and inducing mRNA degradation or inhibiting translation.<sup>11</sup> Currently, over 2,000 known human miRNAs are recorded on the miRBase database,<sup>12</sup> some of which have been characterised as diagnostic biomarkers for diseases such as cardiovascular disease<sup>13</sup> and cancer.<sup>14</sup> Despite several studies postulating the association of various miRNAs with NAFLD, very few have been unequivocally replicated and validated.<sup>15</sup> The lack of consensus may in part be caused by differences in study design, relatively small numbers of cases being studied, technical approaches, and variances within the chosen techniques. The most abundantly expressed miRNA known in adult liver, miR-122-5p, is possibly the strongest candidate NAFLD biomarker, having been consistently corroborated in independent studies.<sup>16–18</sup> More recently, levels of another miRNA, miR-34a-5p, have been shown to increase in serum as NAFLD progresses, and miR-34a-5p has been incorporated in an *in vitro* diagnostic test to assist in the identification of patients with fibrosing steatohepatitis.<sup>19</sup> Our aim was to identify additional serum miRNAs biologically relevant to NAFLD pathophysiology and progression with the capacity to serve as biomarkers. We report a comprehensive and unbiased profile of over 2,000 miRNAs in a large international cohort of patients with histologically characterised NAFLD, with separate discovery and replication performed using diverse techniques in independent patient cohorts.

## Materials and methods

### Patient recruitment and sample collection

The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki. Cases were derived from the European NAFLD Registry (NCT04442334) with the collection and use of NAFLD patient data and samples approved by the relevant ethical committees in the participating centres and all patients having provided informed consent.<sup>20</sup> The collection of blood samples from 'healthy' controls (no reported comorbidities or evidence of NAFLD) was approved by the County Durham and Tees Valley Research Ethics Committee (12/NE/012), and informed consent was given by all donors. Disease stage and grade of NAFLD samples were histologically proven by liver biopsy, with the serum samples taken within 6 months of said biopsy. Liver biopsy specimens were scored centrally by 2 expert liver pathologists (DT and PB) to ensure consistency in the analysis and interpretation of histology data. NAFLD was defined by the presence of steatosis in >5% of hepatocytes. Cases were divided into those with NAFL, if steatosis only or steatosis with lobular inflammation was present, and those with NASH, which is characterised by the presence of both lobular inflammation and hepatocellular ballooning in addition to steatosis.<sup>21,22</sup> NAS was calculated as the unweighted sum of steatosis, ballooning, and lobular inflammation, whereas SAF activity was calculated as the unweighted sum of hepatocyte ballooning and lobular inflammation.<sup>6,7</sup> NAFLD activity was graded as low (NAS 1–4 or SAF activity 0–2) and high (NAS 5–8 or SAF activity 3–4), and fibrosis staged as none/minimal fibrosis stage (F0–F1) and clinically significant (F2–F4).

### Serum sample RNA extraction, library preparation, sequencing, and quality control

Serum samples (15 µl) were processed through the HTG EdgeSeq platform (HTG Molecular Diagnostics, Inc., Tucson, AZ, USA) using the EdgeSeq miRNA Whole Transcriptome Assay according to the manufacturer's instructions. Briefly, the samples were lysed with a 1:1 ratio of HTG biofluids lysis buffer to extract RNA, and the lysates input into nuclease protection assays run on the HTG EdgeSeq processor. Sequence adapters were added to the samples in a 16-cycle PCR. PCR products were purified using AMPure XP beads (Beckman Coulter, High Wycombe, UK) and quantified using a KAPA Library Quantification Kit (KAPA Biosystems, Wilmington, MA, USA). The barcoded libraries were pooled and 2,083 miRNAs (plus 5 negative controls, 1 positive control and 13 housekeeping genes) sequenced at the Genomics Core Facility, Newcastle University, UK, using an Illumina NextSeq 500 System (Illumina Inc., Cambridge, UK) according to the manufacturer's instructions (single end run, 50 bp). FASTQ files were processed by the pre-installed HTG EdgeSeq parser software to align the probe sequences to the results, giving an output of raw count data. Quality control was conducted using the raw counts according to HTG guidelines. Delta mean values within 2 standard deviations (calculated by the manufacturer as 4.8 for serum samples) were considered acceptable.

### miRNA sequencing data analysis

Data for the 2,083 sequenced miRNAs were analysed using RStudio version 1.0.143 (RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, USA, <http://www.rstudio.com/>). Raw

**Table 1. Clinical data for the NAFLD serum samples used in sequencing and quantitative PCR replication.**

	Sequencing cohort (n = 183)	Replication cohort (n = 372)	p value
<b>Patient demographic and clinical characteristics</b>			
Sex (% female)	40%	41%	0.79
Age in years (median [IQR])	54 (44–60)	56 (47–64)	0.007
BMI (median [IQR])	31.6 (27.7–36.1)	31.5 (28.4–36.5)	0.46
AST in IU/L (median [IQR])	37 (27–52)	42 (30–60)	0.014
ALT in IU/L (median [IQR])	53 (36–79)	57 (38–84)	ns
<b>Histological characteristics</b>			
Fibrosis grade			<0.0001
F0	59 (32%)	43 (12%)	
F1	21 (11%)	127 (34%)	
F2	47 (26%)	66 (18%)	
F3	38 (21%)	92 (25%)	
F4	18 (10%)	44 (12%)	
NASH			<0.0001
Yes	128 (70%)	315 (85%)	
No	55 (30%)	57 (15%)	
NAS score >4			0.93
Yes	83 (45%)	171 (46%)	
No	100 (55%)	201 (54%)	
SAF activity score >2			0.52
Yes	67 (37%)	147 (40%)	
No	116 (63%)	225 (60%)	

Values of *p* for continuous or categorical datasets were estimated by the Mann–Whitney test and Chi-square/Fisher’s test, respectively. ALT, alanine aminotransferase; AST, aspartate aminotransferase; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; ns, not significant; SAF, steatosis–activity–fibrosis.

counts were first transformed into counts per million (CPM), and the miRNAs with a CPM ≤100 were filtered out of the raw count dataset. DESeq2 version 1.18.1 (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>)<sup>23</sup> was used to generate a dataset – corrected for batch, age, and centre – from which principal component analysis (PCA) plots were constructed. Outliers were defined by visualisation of the PCA plots and removed from the dataset. Comparisons between each of the histological groupings (control, NAFL, NASH-F0/F1, NASH-F2, NASH-F3, and NASH-F4) and various subgroups were performed using limma version 3.34.9 (<https://bioconductor.org/packages/release/bioc/html/limma.html>).<sup>24</sup> Statistical significance was determined using a Benjamini–Hochberg-adjusted *p* value ≤0.05; no log<sub>2</sub> fold change (logFC) threshold was set. Area under the receiver operating characteristic (AUROC) curves were generated using IBM SPSS Statistics for Windows, version 24 (IBM Corp., Armonk, NY, USA), and significance between 2 AUROC curves was assigned based on the DeLong, DeLong, and Clarke-Pearson method<sup>25</sup> in RStudio version 1.0.143 (RStudio: Integrated development for R. RStudio, Inc., Boston, MA, USA, <http://www.rstudio.com/>).

**miRNA analysis in the replication cohort**

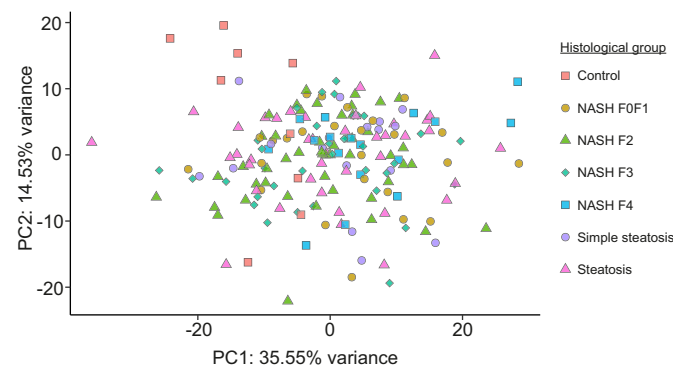
*Total RNA extraction for analysis by quantitative reverse transcriptase PCR*

Total RNA was extracted from serum using the MACHEREY-NAGEL NucleoSpin® miRNA Plasma Kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany; distributed by Fisher Scientific UK, Loughborough, UK) according to the manufacturer’s instructions with only slight modifications to the protocol. Serum (130 µl) was vortexed with 90 µl Buffer MLP and incubated for 3 min at room temperature. As processing controls, 100 fmol of a 5’ phosphorylated exogenous synthetic cel-miR-39-3p and 100 fmol of a non-phosphorylated exogenous synthetic cel-miR-39-

3p were spiked-in (5’-UCACCGGGUGUAAAUCAGCUUG-3’; Integrated DNA Technologies, Inc., Coralville, IA, USA). Total RNA was eluted in 30 µl nuclease-free water and stored at -80°C.

*Complementary DNA synthesis and quantitative reverse transcriptase PCR*

Using the TaqMan™ Advanced miRNA cDNA Synthesis Kit (Thermo Fisher Scientific, Paisley, UK), complementary DNA (cDNA) was synthesised from 3 µl total RNA by reverse transcriptase according to the manufacturer’s instructions. The miR-Amp reaction product was diluted 1:10 in nuclease-free water and stored at -20°C. The expressions of 3 miRNAs, including the phosphorylated spike-in, were quantified using pre-formulated TaqMan™ Advanced miRNA Assays (Thermo Fisher Scientific,



**Fig. 1. Principal component analysis plot for the sequencing data.** Quality controlled, filtered and batch corrected miRNA sequencing data within the plot were coloured according to histological group to aid the identification of clustering by either the first or second principal component. miRNA, micro-RNA; NASH, non-alcoholic steatohepatitis; PC1, principal component 1; PC2, principal component 2.

**Table 2. Differences in miRNA levels detected in the NAFLD-only sequencing analysis.**

miRNA	logFC	logFC SE	Adj. p val.*
<b>NASH grouped relative to NAFL</b>			
miR-193a-5p	0.49	0.11	3.70E-03
miR-378d	0.37	0.09	4.97E-03
miR-378e	0.35	0.09	1.01E-02
miR-4484	0.35	0.09	2.05E-02
<b>NASH with significant fibrosis (F2–F4) relative to NAFL/NASH with minimal fibrosis (F0/F1)</b>			
miR-193a-5p	0.43	0.10	1.79E-03
miR-378e	0.34	0.07	1.79E-03
miR-378d	0.33	0.07	1.85E-03
<b>High NAS (NAS 5–8) relative to low NAS (NAS 1–4)</b>			
miR-193a-5p	0.44	0.09	2.61E-04
miR-378d	0.29	0.07	7.14E-03
miR-320d	0.32	0.09	2.39E-02
miR-378e	0.26	0.07	2.39E-02
miR-320b	0.32	0.09	2.89E-02
miR-320c	0.32	0.09	2.89E-02
miR-320e	0.29	0.08	2.89E-02
<b>High SAF activity (SAF activity 3–4) relative to low SAF activity (SAF activity 0–2)</b>			
miR-193a-5p	0.43	0.11	2.70E-02
miR-378d	0.34	0.09	2.70E-02
miR-378e	0.34	0.09	2.70E-02
miR-19b-3p	-0.74	0.20	2.73E-02

Adj. p val., adjusted p value; logFC, log<sub>2</sub> fold change; miRNA, microRNA; NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; SAF, steatosis-activity-fibrosis.

\* Adjusted p value significance threshold ≤0.05.

**Table 3. miRNAs showing significantly different levels in NASH grouped by fibrosis stage.**

F2			F3			F4		
miRNA	logFC	Adj. p val.	miRNA	logFC	Adj. p val.	miRNA	logFC	Adj. p val.
miR-193a-5p	0.41	4.77E-02	miR-193a-5p	0.72	2.00E-04	miR-193a-5p	0.88	2.32E-03
miR-19b-3p	-0.79	4.77E-02	miR-378d	0.48	4.97E-03	miR-4484	0.63	1.44E-02
miR-3135a	-0.46	4.77E-02	miR-378e	0.43	1.32E-02	miR-378g	0.51	1.79E-02
miR-378d	0.35	4.77E-02	miR-4484	0.45	1.32E-02	miR-6873-3p	-1.24	3.08E-02
miR-378e	0.35	4.77E-02	miR-378a-3p	0.38	1.50E-02			
miR-4534	-0.46	4.77E-02	miR-6780a-5p	0.82	2.59E-02			
miR-649	-0.59	4.77E-02						
miR-670-3p	0.37	4.77E-02						
miR-6769a-5p	0.22	4.77E-02						
miR-6852-5p	0.27	4.77E-02						
miR-25-3p	-0.61	4.83E-02						
miR-330-3p	0.25	4.83E-02						
miR-548ay-5p	-0.51	4.83E-02						
miR-548d-5p	-0.63	4.83E-02						

miRNA levels were assessed by sequencing and are shown relative to NAFL. Adjusted p value threshold ≤0.05. F0/F1 showed no significant differences to NAFL.

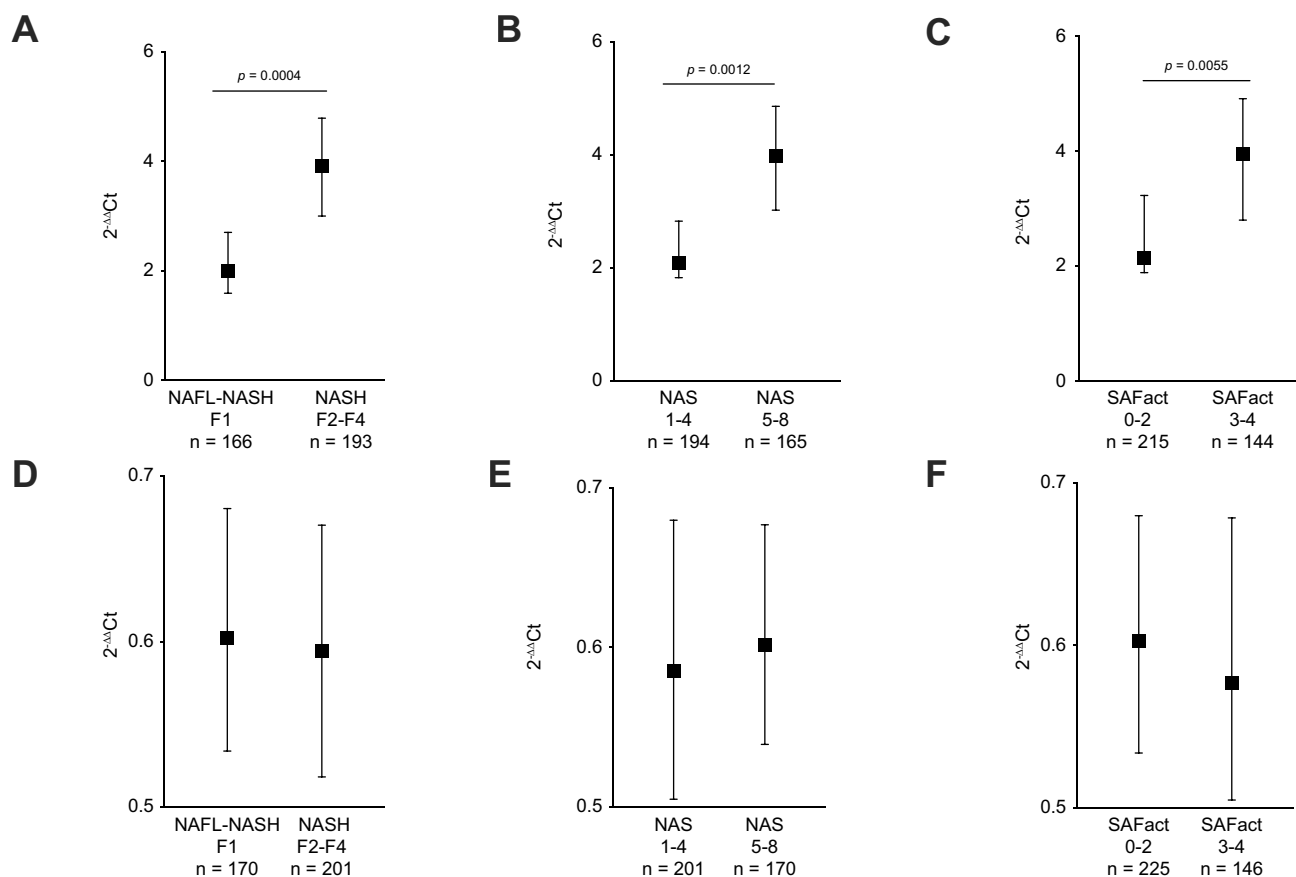
Adj. p val., adjusted p value; logFC, log<sub>2</sub> fold change; miRNA, microRNA; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis.

Paisley, UK): 478293\_mir (for the spiked-in cel-miR-39-3p), 477954\_mir (hsa-miR-193a-5p), and 477855\_mir (hsa-miR-122-5p). Each quantitative PCR (qPCR) mixture contained 5 µl diluted cDNA, 1× TaqMan™ Fast Advanced Master Mix (Thermo Fisher Scientific, Paisley, UK), and 1× TaqMan™ Advanced miRNA Assay, and was made up to 15 µl with nuclease-free water.

TaqMan™ MicroRNA Reverse Transcription Kits (Thermo Fisher Scientific, Paisley, UK) were used to synthesise cDNA from 5 µl total RNA according to the manufacturer's instructions for miRNAs where analysis with material generated via the Advanced miRNA Kit was unsuccessful owing to low levels. The

reverse transcription reaction products were diluted 1:10 in nuclease-free water and stored at -20°C. The expressions of 3 miRNAs, including the non-phosphorylated spike-in, were quantified using pre-formulated standard TaqMan™ MicroRNA Assays (Thermo Fisher Scientific, Paisley, UK): 000200 (cel-miR-39-3p), 464645\_mat (hsa-miR-3687) and 000426 (hsa-miR-34a-5p). Each qPCR reaction mixture contained 2.5 µl diluted cDNA, 1 × TaqMan™ Fast Advanced Master Mix (Thermo Fisher Scientific, Paisley, UK), and 1× TaqMan™ Small RNA Assay (Thermo Fisher Scientific, Paisley, UK) and was made up to 20 µl with nuclease-free water.





**Fig. 2. Replication by qPCR of miR-193a-5p and miR-3687 associations.** Levels of miR-193a-5p are shown for (A) significant fibrosis (NASH F2–F4) relative to minimal fibrosis (NAFL–NASH F0/F1) (n = 359), (B) advanced NAS (NAS 5–8) relative to mild NAS (NAS 1–4) (n = 359), and (C) advanced SAF activity (SAF activity 3–4) relative to mild SAF activity (SAF activity 0–2) (n = 359). Levels of miR-3687 are shown for (D) significant fibrosis (NASH F2–F4) relative to minimal fibrosis (NAFL–NASH F0/F1) (n = 371), (E) advanced NAS (NAS 5–8) relative to mild NAS (NAS 1–4) (n = 371), and (F) advanced SAF activity (SAF activity 3–4) relative to mild SAF activity (SAF activity 0–2) (n = 371).  $2^{-\Delta\Delta C_t}$  was calculated relative to controls for all samples. Median values are shown with 95% CIs. Mann–Whitney *U* tests were performed for all comparisons. NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; qPCR, quantitative PCR; SAF, steatosis–activity–fibrosis.

*qPCR procedure and analysis of replication cohort data*

All qPCRs were run in triplicate in MicroAmp™ Fast Optical 96-Well Reaction Plates (Thermo Fisher Scientific, Paisley, UK) on a QuantStudio™ 5 Real-Time PCR machine (Thermo Fisher Scientific, Paisley, UK). PCR cycling conditions were 50°C for 2 min and 95°C for 20 s, followed by 40 cycles of 95°C for 1 s and 60°C for 20 s. A reference control, made of pooled cDNA, for each assay was run in triplicate on every plate, meaning the raw cycle threshold ( $C_t$ ) values could be corrected for batch effects. The batch corrected data were analysed per miRNA by the  $2^{-\Delta\Delta C_t}$  method.<sup>26</sup> Data were plotted using GraphPad Prism version 8.3.1 (GraphPad Software, San Diego, CA, USA; [www.graphpad.com](http://www.graphpad.com)). Statistical analyses were performed using the Mann–Whitney *U* and Kruskal–Wallis 1-way ANOVA tests.

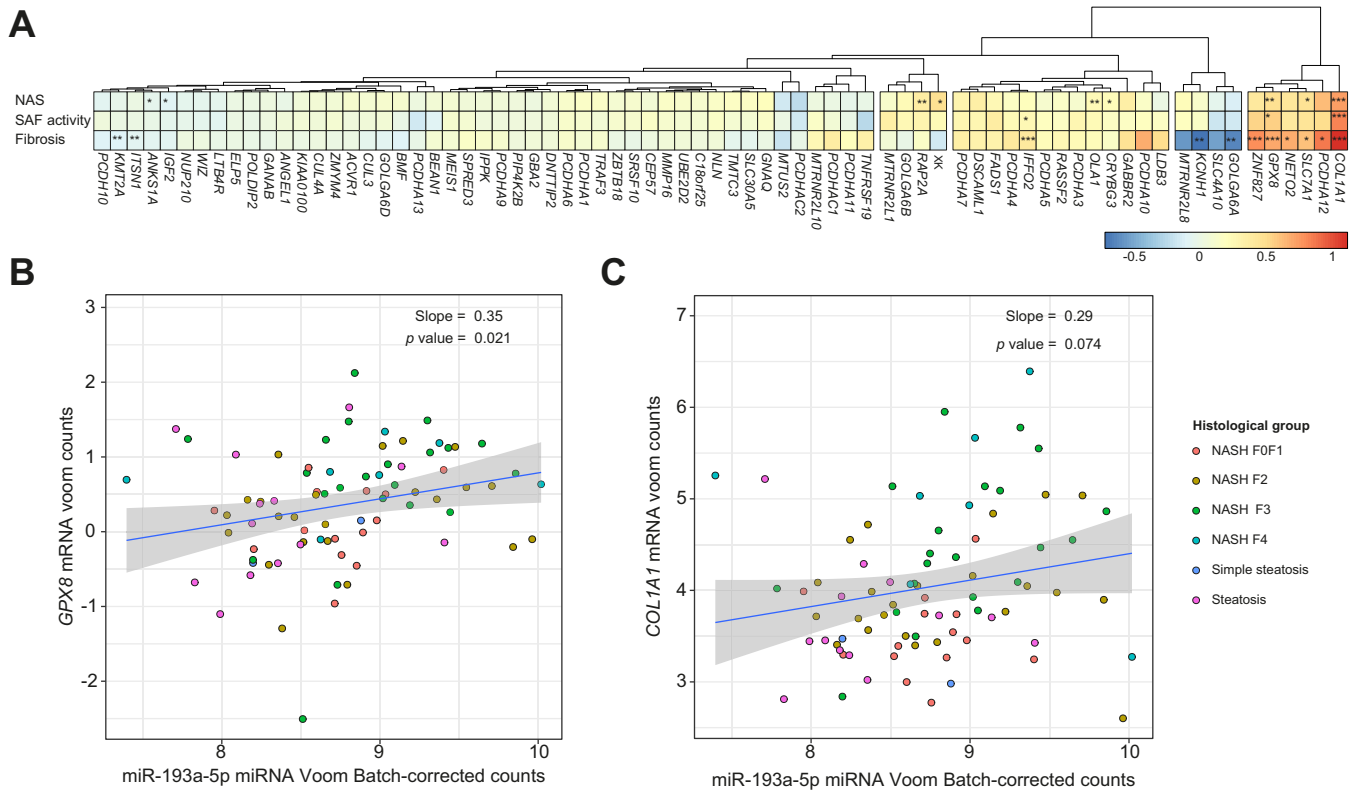
**Target gene predictions and bioinformatic searches**

Two online databases were searched for predicted gene targets of miR-193a-5p, namely, TargetScan<sup>27</sup> and miRDB,<sup>28</sup> and the overlapping genes between the 2 were identified. Common genes were identified between the resulting predicted genes and the differentially expressed genes in NAFLD liver as described by

Govaere *et al.*<sup>29</sup> Normalised and transformed counts of miR-193a-5p, *GPX8*, and *COL1A1* were correlated for the 80 overlapping samples for which miRNA-seq and RNA-seq data were available; a linear model was performed to obtain the slopes and *p* values of the correlations. Single-cell RNA-seq data for *GPX8* and *COL1A1* were extracted from the Liver Cell Atlas gene browser on 13 August 2020.<sup>30</sup> Data on gene expression were obtained from the Gene Page of the Genotype-Tissue Expression (GTEx) Portal (GTEx Analysis Release V8; dbGaP Accession phs000424.v8.p2).

**Hep G2 cell culture and treatment**

The human hepatoblastoma Hep G2 cell line (ATCC® HB-8065™, Middlesex, UK), was cultured as described previously.<sup>29</sup> To assess effects on miRNA expression, cells were treated for 24 h with the following: (i) 1% bovine serum albumin-conjugated palmitic acid (250 μM), (ii) 1% bovine serum albumin-conjugated oleic acid (500 μM), (iii) 1% bovine serum albumin-conjugated palmitic and oleic acid (250/500 μM) combined solution, and (iv) 1% bovine serum albumin as a lipid-loading control. All treatments were performed in



**Fig. 3. Differentially expressed predicted target genes of miR-193a-5p in liver RNA-seq.** (A) RNA-seq data from liver tissue were analysed for 3 comparisons: advanced NAS (NAS 5–8) relative to mild NAS (NAS 1–4), advanced SAF activity (SAF activity 3–4) relative to mild SAF activity (SAF activity 0–2), and significant fibrosis (NASH F2–F4) relative to minimal fibrosis (NAFL–NASH F0/F1). The genes that overlapped with those predicted by *in silico* tools to be targets of miR-193a-5p are shown in the heatmap. Hierarchical clustering is based on levels of fold change in gene expression in the liver tissue. Statistical significance in the RNA-seq data was determined using a Benjamini–Hochberg adjusted  $p$  value  $\leq 0.05$ : \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , and \*\*\* $p \leq 0.001$ . (B and C) Linear models of 80 overlapping NAFLD samples between the miRNA-seq and RNA-seq datasets for miR-193a-5p and (B) GPX8 and (C) COL1A1. AFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; SAF, steatosis–activity–fibrosis.

triplicate. Total RNA was isolated from the cells as described previously,<sup>29</sup> and cDNA was synthesised and qPCR performed as described above using the TaqMan™ Advanced miRNA cDNA Synthesis Kit and TaqMan™ Advanced miRNA Assay for miR-193a-5p.

## Results

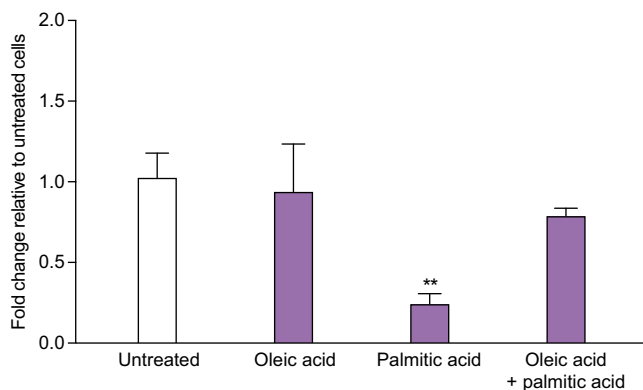
### Study cohort description

The current study comprised cases that were histologically characterised and represented the full spectrum of NAFLD severity from NAFL to NASH F4 (cirrhosis). The samples were divided between discovery ( $n = 183$ ) and replication ( $n = 372$ ) cohorts. Discovery cases to cover the range of NAFLD phenotypes were selected from those already recruited within the European NAFLD Registry<sup>20</sup> by July 2018 with replication cases selected from the larger cohort available by June 2019. Demographic and clinical data are reported in Table 1. There were some significant differences between the cohorts. In particular, the discovery cohort included a relatively large number of NAFL cases without detectable fibrosis or inflammation to allow detailed comparisons between early and late disease. The replication cohort was

significantly older with more severe disease and included a wider spectrum of phenotypes. To obtain baseline measurements for comparison with cases with NAFLD, 10 individuals from a healthy control group (3 male, 7 female, mean age 31 years) were included in the analysis with the discovery cohort. A further 15 control individuals (9 male, 6 female, mean age 42 years) were included in the replication cohort analyses.

### Comparison of NAFLD serum miRNA profiles with controls

The serum miRNA profiles of 183 patients across the NAFLD spectrum and 10 population controls were generated using the HTG EdgeSeq and Illumina NextSeq technologies. Complete sequence data are available as GSE185062 on Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). Initial PCA confirmed that there was no clustering by centre, plate/batch, or age, indicating that batch correction was effective (Fig. S1). PCA showed that the control samples appeared to cluster together, suggesting they had a different miRNA profile to the NAFLD samples, although some overlap was observed (Fig. 1). Relative to controls, levels of 275 serum miRNAs were different in NAFLD patients, 165 increasing and 110 decreasing (Table S1). The most significantly different miRNA was miR-



**Fig. 4. In vitro functional assessment of miR-193a-5p expression in Hep G2 cells.** Hep G2 cells were treated with fatty acids (oleic acid [500  $\mu$ M], palmitic acid [250  $\mu$ M], or a combination of oleic [500  $\mu$ M] and palmitic acid [250  $\mu$ M]) for 24 h. All treatments were performed in triplicate, and qPCRs of each were performed in triplicate. Data were normalised to the control condition (untreated) using the  $2^{-\Delta\Delta Ct}$  method. An unpaired Student's *t*-test was performed for all conditions relative to the control conditions (\*\**p*  $\leq$  0.01). Data are presented as the mean with error bars representing the SEM. qPCR, quantitative PCR.

3687 for which decreased levels in NAFLD were seen (fold change [FC] 0.10, *p* =  $4.57 \times 10^{-28}$ ) (Table S2) When cases with NASH and different stages of fibrosis were compared with the population controls, in excess of 200 significant differences in miRNA levels were detected for each fibrosis stage, with the top 200 for each listed in Table S2. The most significant alteration was again the decrease seen for miR-3687 but for NASH F0/F1, F2, and F3 only.

#### Comparison of NAFLD miRNA profiles with NAFL as baseline

We next conducted an analysis amongst NAFLD cases only, excluding population controls. Using NAFL as the baseline, we initially looked for serum miRNA changes in NASH cases generally; 4 miRNAs showed increased levels with no miRNAs decreased (Table 2). The miRNA showing the largest increase was miR-193a-5p (FC 1.34, *p* =  $3.70 \times 10^{-03}$ ). We subsequently characterised the miRNA profiles of NASH cases with particular stages of fibrosis (F0–F4) against NAFL (Table 3). The miRNA profile for NASH-F0/F1 was similar to that for NAFL; thereafter, progressive changes in miRNA levels were seen for cases with NASH-F2, NASH-F3, and NASH-F4 (cirrhosis), suggesting that it was the stage of fibrosis rather than grade of steatohepatitic activity that was driving the association. All changes seen in F3 and F4 were increases in miRNA levels, but for F2, both increases and decreases in levels were detected. MiR-193a-5p showed the most significant change with a progressive increase in levels from F2 (FC 1.32) via F3 (FC 1.65) to F4 (FC 1.84) (Table 3). These increases in miR-193a-5p levels with increasing fibrosis score were also in line with those seen for the NAFLD cases compared with population controls (Table S2).

To assess the potential for distinguishing between mild and more clinically significant disease, we undertook subgroup comparisons of the miRNA profile in relation to disease activity and fibrosis stage. Three miRNAs showed increased levels in clinically significant fibrosis compared with minimal disease (Table 2); the most statistically significant of these was miR-193a-5p (FC 1.35, *p* =  $1.79 \times 10^{-03}$ ). When the grade of

steatohepatitis was considered using NAS and SAF activity, 7 and 4 miRNAs respectively showed altered levels, with all but 1 of these changes an increase with increasing disease severity (Table 2). MiR-193a-5p again showed the most significant increase (NAS FC = 1.36, *p* =  $2.61 \times 10^{-04}$ ; SAF activity FC = 1.35, *p* =  $2.70 \times 10^{-02}$ ).

In light of its association with disease severity, we evaluated miR-193a-5p as a potential circulating biomarker of disease stage and activity by assessing its diagnostic performance in the sequencing dataset using AUROC statistics. AUROCs were also calculated and plotted for miR-122-5p as well as simple composite scores (the aspartate aminotransferase/alanine aminotransferase [AST/ALT] ratio and the fibrosis-4 [FIB-4] score).<sup>31</sup> Relevant curves and values are detailed in Fig. S2. The AUROC for miR-193a-5p was 0.92 for discriminating NAFLD from population controls, whereas the values for differentiating mild disease from significant fibrosis (F0–1 vs. F2–4), high NAS, and high SAF activity were 0.68, 0.73, and 0.71, respectively. MiR-193a-5p significantly outperformed AST/ALT ratio when discriminating between mild vs. significant fibrosis (*p* = 0.04), high NAS (*p* =  $6.00 \times 10^{-05}$ ), and high SAF activity (*p* = 0.03). It also had better discriminatory ability than FIB-4 as an indicator of high NAS (*p* =  $3.40 \times 10^{-03}$ ) and exhibited comparable performance with FIB4 for fibrosis. The discriminatory ability of miR-193a-5p was better than that of miR-122-5p for fibrosis (*p* = 0.019), NAS (*p* = 0.012), and SAF activity (*p* =  $1.53 \times 10^{-03}$ ).

#### Replication of miR193a-5 and previously reported miRNAs as predictors of NAFLD progression using qPCR in an independent cohort of samples

We sought to replicate selected findings from our sequencing studies relating to miR-193-5p and miR-3687 and additionally quantify serum levels of 2 miRNAs (miR-122-5p and miR-34a-5p) that have previously been reported as relevant to NAFLD,<sup>16–19</sup> using the independent replication cohort of 372 NAFLD cases; healthy control samples (*n* = 15) were also included. Limited volumes of serum and inability to analyse using the Advanced miRNA Kit prevented replication of other interesting miRNA signals. Consistent with the results of our sequencing analysis, miR-193a-5p showed significantly higher levels in NAFLD cases with greater disease severity (*i.e.* more significant fibrosis and high NAS/SAF activity) (Fig. 2). However, the differences observed in miR-3687 levels did not vary significantly with grade or stage of disease in the replication cohort (Fig. 2). Levels of miR-34a-5p were also significantly increased with a higher grade of NAFLD activity (by both NAS and SAF activity) and with more significant fibrosis (Fig. S3). In contrast, miR-122-5p levels did not show significance by fibrosis stage, but levels increased significantly in higher grades of NAFLD activity by both NAS and SAF activity (Fig. S3).

#### Correlation of miR-193a-5p target gene predictions with differential hepatic gene expression measured by RNA-seq in NAFLD

In view of the significant increase in serum miR-193a-5p levels seen as disease activity increased and its strong discriminatory ability, we sought to investigate the functional significance of miR-193a-5p in disease progression. Bioinformatic analysis using TargetScan<sup>27</sup> and miRDB<sup>28</sup> predicted that 204 and 327 genes, respectively, were targets of miR-193a-5p, with an overlap of 78 genes between both prediction sets. To better examine the transcriptional correlates of miR-193a-5p, we reanalysed our



recently reported hepatic transcriptomics dataset from 206 histologically characterised NAFLD cases to identify genes differentially expressed between mild and severe steatohepatitis and/or fibrosis.<sup>29</sup> We then projected the list of 78 *a priori* predicted miR-193a-5p targets onto these data. The resulting differentially expressed genes for each comparison are detailed in Fig. 3. Nine predicted gene targets of miR-193a-5p were significantly differentially expressed in the liver of patients with greater steatohepatic activity by NAS (*COL1A1*, *CRYBG3*, *GPX8*, *OLA1*, *RAP2A*, *SLC7A1*, *XK*, *ANKS1A*, and *IGF2*) and 3 by SAF activity (*COL1A1*, *GPX8*, and *IFFO2*), and 11 gene targets were differentially expressed in the liver of patients with significant fibrosis (*COL1A1*, *GPX8*, *IFFO2*, *NETO2*, *PCDHA12*, *SLC7A1*, *ZNF827*, *GOLGA6A*, *ITSN1*, *KCNH1*, and *KMT2A*). Additionally, the correlation between serum miR-193a-5p and levels in the liver of the 2 genes that were differentially expressed in all 3 comparisons, *COL1A1* and *GPX8*, was investigated in a subgroup of 80 samples for which both serum miRNA-seq and liver RNA-seq data were available (Fig. 3). A significant positive correlation between miR-193a-5p and *GPX8* was observed (slope = 0.35,  $p = 0.021$ ), and there was a similar trend with *COL1A1* (slope = 0.29,  $p = 0.074$ ). Moreover, single-cell RNA-seq data extracted from the Liver Cell Atlas gene browser<sup>30</sup> showed that both *GPX8* and *COL1A1* followed a similar pattern of increased expression in cirrhotic compared with uninjured tissue in the mesenchyme (Fig. S4).

#### Quantification of miR-193a-5p expression in a hepatoma cell line

We utilised data from the GTEx Portal to visualise the tissue-specific expression of miR-193a-5p, confirming that – in addition to the renal medulla, omentum and subcutaneous adipose tissue, mammary tissue, and cultured fibroblasts – miR-193a-5p was expressed in the liver (Fig. S5). Considering this evidence, we characterised the expression of miR-193a-5p *in vitro* in a model system of human hepatocytes in the context of lipid accumulation as seen in NAFLD (Fig. 4). Hep G2 cells were loaded with fatty acids, and miR-193a-5p expression was quantified by qPCR. After a 24-h treatment with lipotoxic palmitic acid, miR-193a-5p expression was decreased significantly relative to the control condition ( $p = 0.009$ ). A less lipotoxic but more steatogenic lipid, oleic acid,<sup>32</sup> had a negligible effect relative to the control condition, and, also in line with a previous report,<sup>32</sup> there was no effect when the 2 lipids were combined.

#### Discussion

An unbiased profile of circulating miRNAs in NAFLD patients and controls has been generated using high-throughput sequencing techniques. Although other studies have identified some changes in serum miRNA levels in NAFLD,<sup>17,33,34</sup> our study is the largest of its kind – in terms of sample number and number of miRNAs analysed – spanning the full spectrum of the disease including cirrhosis. Although targeting only known miRNAs and unable to detect novel miRNAs, within these limitations we conducted a comprehensive and unbiased profile of miRNAs in NAFLD. Our sequencing approach revealed hundreds of miRNAs showing different levels at every NAFLD fibrosis stage relative to the controls, suggesting that these levels might correlate with, and therefore potentially predict, disease progression. We also saw

numerous significant differences in miRNA levels within the entire NAFLD group compared with controls. Altered serum miRNA levels could be a useful diagnostic tool for NAFLD generally, but identifying such changes was not an aim of this study.

By combining the NAFLD sequencing data with an independent qPCR replication, we have shown consistently that serum levels of miR-193a-5p can distinguish between mild and advanced disease activity and fibrosis. Our finding for miR-193a-5p is in line with a recent report that performed miRNA sequencing of plasma samples from NAFLD cases but focussed on predominantly mild steatosis and fibrosis.<sup>34</sup> We also saw increased levels of miR-122-5p, which was generally in line with previous reports,<sup>16,17,35</sup> apart from a failure to detect significant differences with increasing fibrosis severity. This suggests that miR-193a-5p is a more sensitive biomarker of NAFLD progression than miR-122-5p. In addition to miR-193a-5p and miR-122-5p, we detected several other miRNAs that show differences in levels as NAFLD progresses, although the effect was not as significant as that for miR-193a-5p and overall levels of these miRNAs were lower. These additional significant miRNAs included miR-378d and miR-378e, which both increased in NASH and advanced fibrosis. This finding is in line with a recent sequencing study by Zhang *et al.*,<sup>34</sup> who, while reporting that these 2 miRNAs were present at higher levels in plasma from NAFLD cases compared with controls and showed a significant relationship with extent of fibrosis, also found lower levels in plasma compared with miR-193a-5p, in both NAFLD cases and controls. We also saw significantly raised levels of miR-4484 in NASH compared with NAFL. The previous sequencing study reported levels of this miRNA increased in severe steatosis but did not include data on inflammation.<sup>34</sup> Our sequencing analysis did not detect changes in levels of miR-34a, another previously reported miRNA that increased in NAFLD,<sup>19</sup> but we were able to confirm previous data for this relatively low-abundance miRNA as a significant marker for NAFLD progression by use of a highly sensitive qPCR assay. Indicating its potential usefulness as a biomarker, serum miR-193a-5p achieved high AUROC values in our study. There is increasing evidence that at least a proportion of circulating miRNAs is encapsulated within extracellular vesicles in NAFLD.<sup>36</sup> The sequencing approach we used detects all miRNAs within serum without distinguishing between those present in vesicles and those circulating outside vesicles that are possibly protein-bound.

There is evidence from other studies that levels of miR-193a-5p are decreased in liver tissue in fibrosis<sup>37</sup> and in hepatocellular carcinoma<sup>38</sup> and that hepatic miR-193a-5p levels exhibit an inverse correlation with serum levels<sup>37</sup> similar to that reported for other miRNAs including miR-122-5p.<sup>17,39</sup> This would also be broadly consistent with our observations in Hep G2 cells where proapoptotic palmitic acid<sup>32</sup> decreases miR-193a levels but steatogenic oleic acid, either alone or in combination with palmitic acid, has no significant effect. The miR-193a target genes, found to show altered expression during NAFLD in our transcriptomic analysis, all increased in expression, which is typical of genes regulated by miRNA with a decreased miRNA level facilitating increased mRNA levels. In a previous study, miR-193a-5p has been shown to act within a network of miRNAs to modify the transforming growth factor-

beta (TGF- $\beta$ )-dependent regulation of extracellular matrix-related genes in hepatic fibrosis.<sup>37</sup> It is likely, therefore, that miR-193a-5p in NAFLD will exert its effects on target genes as part of a large interaction network over a longer time course rather than, for example, as an immediate response to an acute stress. One of the other miRNAs that showed increased serum levels in severe fibrosis, miR-378d, has also been demonstrated to show decreased expression during fibrosis in an animal model.<sup>40</sup>

Significantly differentially expressed predicted gene targets of miR-193a-5p in our transcriptomics data include *RAP2A*, *IGF2*, and *SLC7A1*, which have all been associated with hepatic function, liver disease, and/or steatosis.<sup>41–43</sup> Most notably, we found *COL1A1* and *GPX8* showed increased expression by all 3 comparisons. A separate comparison using single-cell RNA-seq data also showed increased expression in cirrhotic cells compared with uninjured mesenchyme cells.<sup>30</sup> *COL1A1* encodes an extracellular matrix component that is extensively deposited in the development of scarring and fibrosis.<sup>44</sup> Accordingly, it is a highly biologically plausible and relevant target of miR-193a-5p and has previously been identified as part of a transcriptomic gene signature associated with advanced NAFLD.<sup>29</sup> MiR-193a-5p has also been shown to interact with *COL1A1* in colorectal cancer cells and is known to be involved in extracellular matrix deposition.<sup>37,45</sup> It is therefore plausible that secretion of miR-193a-5p by hepatocytes permits an increase in *COL1A1* expression in stellate cells and thus contributes to hepatic fibrosis as proposed by others.<sup>46</sup>

*GPX8* is an endoplasmic reticulum (ER)-resident member of the glutathione peroxidase family that functions to protect cells from oxidative damage<sup>47</sup>; oxidative damage may in turn lead to ER stress and is associated with metabolically driven NAFLD pathologies.<sup>48</sup> Roy *et al.*<sup>37</sup> have previously described the expression of miR-193a-5p in the liver and its role as a member of a network of miRNAs modifying the TGF- $\beta$ -dependent regulation of extracellular matrix-related genes expression in hepatic stellate cells to modify the balance between deposition and resorption of liver fibrosis. Recently, it has been demonstrated that *GPX8* expression increases during epithelial-to-mesenchymal transition in several cell lines with evidence that increased *GPX8* affects IL-6/JAK-STAT3 signalling via the soluble IL6 receptor.<sup>49</sup> Given that hepatic miR-193a-5p expression is known to exhibit an inverse correlation with serum levels,<sup>37</sup> we therefore postulate that the increase in serum miR-193a-5p we observed in patients with advanced NAFLD is indicative of increased hepatocyte export. This could induce increased *GPX8* expression in other hepatic cells as they respond to oxidative stress and undergo further transitions during NAFLD

progression. Indeed, we observed a significant positive correlation between serum miR-193a-5p and hepatic *GPX8*. Nonetheless, further investigations are required to identify the specific cells in which *GPX8* acts, as low hepatocyte expression in the single-cell RNA-seq dataset suggests that it may exert its effects elsewhere in the liver.<sup>30</sup>

Published reports on miRNAs in NAFLD are seldom in complete agreement,<sup>50</sup> and as such, one should be mindful of pertinent limitations. Discordance between studies could be caused by sampling and interpretation errors, and so the histological classification of patients' disease might be inconsistent from the outset. Despite being a component of an *in vitro* diagnostic test for NASH,<sup>19</sup> our sequencing analysis filtered out miR-34a-5p because of low levels, yet subsequent qPCR was in fact able to detect it. Conversely, the strongest signal in NAFLD relative to controls, miR-3687, was not confirmed by qPCR. These examples highlight that sequencing should always be replicated independently using a different methodology to test the strength of the original dataset. This approach combined with use of a replication cohort enabled us to confirm the novel miR-193a-5p signal. However, the qPCR approach also has limitations, and we were unable to replicate all the associations found by sequencing owing to limited serum availability. A higher throughput approach for replication, possibly using nanoString or a custom microarray would be helpful if a further replication cohort was available. Further studies on the functional significance of miR-193a-5p, especially direct confirmation of effects on specific target genes, and the identification of cells and possible extra-hepatic tissues in which these interactions and downstream effects occur are also still needed.

Overall, we have obtained an unbiased global profile of circulating miRNAs in NAFLD, and our study has identified a number of miRNAs that display changes in levels during disease progression. In particular, the increase in circulating miR-193a-5p was replicated independently, and we showed that several functionally relevant gene targets of this miRNA were differentially expressed in liver tissue. We have also further investigated the reported association of miR-122-5p and miR-34a-5p with NAFLD, finding that miR-122-5p was not a significant predictor of fibrosis progression but confirming that miR-34a-5p is a sensitive, although difficult to detect, marker. MiR-193a-5p had a superior ability to discriminate between mild and advanced NASH compared with AST/ALT and FIB-4 and could have value as an additional biomarker for progression, possibly in combination with other biologically relevant indices. The current study demonstrates how global miRNA profiling may provide both new insights into disease pathophysiology and identify promising new liver disease biomarkers.

## Abbreviations

ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic; cDNA, complementary DNA; CPM, counts per million; C<sub>t</sub>, cycle threshold; ER, endoplasmic reticulum; FC, fold change; FIB-4, fibrosis-4; FLIP, fatty liver inhibition of progression; GTEx, Genotype-Tissue Expression; log<sub>2</sub>FC, log<sub>2</sub> fold change; miRNA, microRNA; NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; PCA, principal component analysis; qPCR, quantitative PCR; SAF, steatosis-activity-fibrosis; TGF- $\beta$ , transforming growth factor-beta.

## Financial support

This study has been performed as part of the LITMUS (Liver Investigation: Testing Marker Utility in Steatohepatitis) project, which has received funding from the Innovative Medicines Initiative (IMI2) Program of the European Union under Grant Agreement 777377; this Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. The current study was also supported by the Newcastle NIHR Biomedical Research Centre and the European NAFLD Registry.

**Conflicts of interest**

The authors have no potential conflicts (financial, professional, or personal) directly relevant to the manuscript. Please refer to the accompanying ICMJE disclosure forms for further details.

**Authors' contributions**

Conceptualisation: KJ, OG, QMA, AKD. Methodology: KJ, PJJ, OG. Software: PJJ, SJC. Validation: KJ. Formal analysis: KJ, PJJ. Investigation: KJ, PJJ, OG, MJB, SHC, DT, MZ, PB. Resources: JC, ME, GA, KC, JMS, JB, VR, EB, QMA, AKD; Data curation: KJ, PJJ. Writing – original draft: KJ, QMA, AKD; Writing – review and editing: KJ, PJJ, OG, MJB, SHC, SJC, DT, PB, MJB, JFC, ME, GA, KC, JMS, JB, VR, EB, QMA, AKD. Project supervision: QMA, AKD. Funding acquisition: QMA, AKD

**Data availability statement**

The complete miRNA sequence data are available on Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) (GSE185062). Other data from the study are available from the corresponding authors upon reasonable request.

**LITMUS Consortium Investigators**

Newcastle University, UK: Quentin M. Anstee, Pierre Bedossa, James Clark, Simon Cockell, Heather J. Cordell, Ann K. Daly, Rebecca Darlay, Christopher P. Day, Olivier Govaere, Tim Hardy, Katherine Johnson, Yang-Lin Liu, Fiona Oakley, Jeremy Palmer, Rachel Queen, Dina Tiniakos, Kristy Wonders, Michalina Zatorska

Academisch Medisch Centrum bij de Universiteit van Amsterdam, Netherlands: Patrick M. Bossuyt, Adriaan G. Holleboom, Hadi Zafarmand, Yasaman Vali, Jenny Lee

Hôpital Beaujon, Assistance Publique: Hôpitaux de Paris, France and Institute of Cardiometabolism and Nutrition, Paris, France: Karine Clement, Raluca Pais, Vlad Ratziu

University Medical Centre Mainz, Germany: Jörn M. Schattenberg, Detlef Schuppan

University of Cambridge, UK: Michael Allison, Sergio Rodriguez Cuenca, Vanessa Pellegrinelli, Michele Vacca, Antonio Vidal-Puig

Örebro University, Sweden: Tuulia Hyötyläinen, Aidan McGlinchey, Matej Orešič, Partho Sen

Centre for Cooperative Research in Biosciences, Spain: Jose Mato, Óscar Millet

University of Bern, Switzerland: Jean-Francois Dufour

University of Oxford, UK: Jeremy F. Cobbald, Stephen Harrison, Stefan Neubauer, Michael Pavlides, Ferenc Mozes, Salma Akhtar

Perspectum Diagnostics Ltd, UK: Rajarshi Banerjee, Matt Kelly, Elizabeth Shumbayawonda, Andrea Dennis, Charlotte Ercicum

Servicio Andaluz de Salud, Spain: Manuel Romero-Gomez, Rocío Gallego-Durán, Isabel Fernández

Nordic Bioscience A/S, Denmark: Morten Karsdal, Diana Leeming, Mette Juul Fisker, Elisabeth Erhardtsen, Daniel Rasmussen, Per Qvist, Antonia Sinesi

Integrated BioBank of Luxembourg Foundation, Luxembourg: Estelle Sandt, Maria Manuela Tonini

University of Torino, Italy: Elisabetta Bugianesi, Maurizio Parola, Chiara Rosso

University of Firenze, Italy: Fabio Marra

Consiglio Nazionale Delle Ricerche, Pisa, Italy: Amalia Gastaldelli

University of Angers, France: Jerome Boursier

Antwerp University Hospital, Belgium: Sven Francque

Linköping University, Sweden: Mattias Ekstedt, Stergios Kechagias

University of Helsinki, Finland: Hannele Yki-Järvinen, Kimmo Porthan

University Medical Centre Utrecht, Netherlands: Saskia van Mil

National & Kapodistrian University of Athens, Greece: George Papatheodoridis

Faculdade de Medicina da Universidade de Lisboa, Portugal: Helena Cortez-Pinto

Università degli Studi di Milano, Italy: Luca Valenti

University of Palermo, Italy: Salvatore Petta

Università Cattolica del Sacro Cuore, Italy: Luca Miele

University Hospital Würzburg, Germany: Andreas Geier

Universitätsklinikum Aachen University, Germany: Christian Trautwein

University of Nottingham, UK: Guruprasad P. Aithal

Antaros Medical AB, Sweden: Paul Hockings

University of Birmingham, UK: Phil Newsome

iXscient Ltd, UK: David Wenn

Faculdade de Farmácia da Universidade de Lisboa, Lisbon, Portugal: Cecília Maria Pereira Rodrigues

Genfit SA, France: Rémy Hanf, Pierre Chaumat, Christian Rosenquist

Intercept Pharma Europe Ltd, UK: Aldo Trylesinski

One Way Liver S.L., Spain: Pablo Ortiz

Eli Lilly and Company Ltd, USA: Kevin Duffin

Pfizer Ltd, USA: Carla Yunis, Melissa Miller, M. Julia Brosnan, Theresa Tuthill

Boehringer-Ingelheim International GmbH, Germany: Judith Ertle, Ramy Younes

SomaLogic Inc., USA: Leigh Alexander, Rachel Ostroff

Novo Nordisk A/S, Denmark: Mette Skalshøj Kjær

Ellegaard Göttingen Minipigs AS, Denmark: Lars Friis Mikkelsen

Novartis Pharma AG, Switzerland: Clifford Brass, Lori Jennings, Maria-Magdalena Balp, Miljen Martić

Takeda Pharmaceuticals International GmbH, Switzerland: Guido Hanauer

AstraZeneca AB, Sweden: Sudha Shankar, Richard Torstenson

Echosens, France: Céline Fournier

Resoundant Inc., USA: Richard Ehman, Michael Kalutkiewicz, Kay Pepin

Bristol-Myers Squibb Company Corp., USA: Joel Myers, Diane Shevell

HistoIndex Pte Ltd, Singapore: Gideon Ho

Allergan: Henrik Landgren

Gilead Science Inc., USA: Rob Myers

Research Triangle Institute: Lynda Doward, Diane Whalley, James Twiss

**Supplementary data**

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhepr.2021.100409>.

**References**

- [1] Younossi ZM. Non-alcoholic fatty liver disease – a global public health perspective. *J Hepatol* 2019;70:531–544.
- [2] Anstee QM, Day CP. The genetics of NAFLD. *Nat Rev Gastroenterol Hepatol* 2013;10:645–655.
- [3] Anstee QM, Reeves HL, Kotsiliti E, Govaere O, Heikenwalder M. From NASH to HCC: current concepts and future challenges. *Nat Rev Gastroenterol Hepatol* 2019;16:411–428.
- [4] Anstee QM, Darlay R, Cockell S, Meroni M, Govaere O, Tiniakos D, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort. *J Hepatol* 2020;73:505–515.
- [5] Anstee QM, Seth D, Day CP. Genetic factors that affect risk of alcoholic and nonalcoholic fatty liver disease. *Gastroenterology* 2016;150:1728–1744 e1727.
- [6] 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.
- [7] Bedossa P, FLIP Pathology Consortium. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. *Hepatology* 2014;60:565–575.
- [8] 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.
- [9] Rinella ME, Tacke F, Sanyal AJ, Anstee QM, participants of the A-EW. Report on the AASLD/EASL joint workshop on clinical trial endpoints in NAFLD. *J Hepatol* 2019;71:823–833.
- [10] Pillai RS. MicroRNA function: multiple mechanisms for a tiny RNA? *RNA* 2005;11:1753–1761.
- [11] O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front Endocrinol (Lausanne)* 2018;9:402.
- [12] Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 2019;47:D155–D162.
- [13] Zhou SS, Jin JP, Wang JQ, Zhang ZG, Freedman JH, Zheng Y, et al. miRNAs in cardiovascular diseases: potential biomarkers, therapeutic targets and challenges. *Acta Pharmacol Sin* 2018;39:1073–1084.
- [14] Wang H, Peng R, Wang J, Qin Z, Xue L. Circulating microRNAs as potential cancer biomarkers: the advantage and disadvantage. *Clin Epigenetics* 2018;10:59.



- [15] López-Riera M, Conde I, Quintas G, Pedrola L, Zaragoza A, Perez-Rojas J, et al. Non-invasive prediction of NAFLD severity: a comprehensive, independent validation of previously postulated serum microRNA biomarkers. *Sci Rep* 2018;8:10606.
- [16] Dongiovanni P, Meroni M, Longo M, Fargion S, Fracanzani AL. miRNA signature in NAFLD: a turning point for a non-invasive diagnosis. *Int J Mol Sci* 2018;19.
- [17] Pirola CJ, Fernandez Gianotti T, Castaño GO, Mallardi P, San Martino J, Mora Gonzalez Lopez Ledesma M, et al. Circulating microRNA signature in non-alcoholic fatty liver disease: from serum non-coding RNAs to liver histology and disease pathogenesis. *Gut* 2015;64:800–812.
- [18] Cheung O, Puri P, Eicken C, Contos MJ, Mirshahi F, Maher JW, et al. Nonalcoholic steatohepatitis is associated with altered hepatic MicroRNA expression. *Hepatology* 2008;48:1810–1820.
- [19] Harrison SA, Ratzliff V, Boursier J, Francque S, Bedossa P, Majd Z, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol* 2020;5:970–985.
- [20] Hardy T, Wonders K, Younes R, Aithal GP, Aller R, Allison M, et al. The European NAFLD Registry: a real-world longitudinal cohort study of nonalcoholic fatty liver disease. *Contemp Clin Trials* 2020;98:106175.
- [21] Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology* 2012;56:1751–1759.
- [22] Sanyal AJ, Brunt EM, Kleiner DE, Kowdley KV, Chalasani N, Lavine JE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology* 2011;54:344–353.
- [23] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- [24] 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.
- [25] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.
- [26] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  method. *Methods* 2001;25:402–408.
- [27] Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 2015;4:e05005.
- [28] Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res* 2020;48:D127–D131.
- [29] Govaere O, Cockell S, Tiniakos D, Queen R, Younes R, Vacca M, et al. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci Transl Med* 2020;12:eaba4448.
- [30] Ramachandran P, Dobie R, Wilson-Kanamori JR, Dora EF, Henderson BEP, Luu NT, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* 2019;575:512–518.
- [31] McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut* 2010; 59:1265–1269.
- [32] Ricchi M, Odoardi MR, Carulli L, Anzivino C, Ballestri S, Pinetti A, et al. Differential effect of oleic and palmitic acid on lipid accumulation and apoptosis in cultured hepatocytes. *J Gastroenterol Hepatol* 2009;24:830–840.
- [33] Tan Y, Ge G, Pan T, Wen D, Gan J. A pilot study of serum microRNAs panel as potential biomarkers for diagnosis of nonalcoholic fatty liver disease. *PLoS One* 2014;9:e105192.
- [34] Zhang X, Mens MMJ, Abozaid YJ, Bos D, Darwish Murad S, de Kneegt RJ, et al. Circulatory microRNAs as potential biomarkers for fatty liver disease: the Rotterdam study. *Aliment Pharmacol Ther* 2021;53:432–442.
- [35] Jampoka K, Muangpaisarn P, Khongnomnan K, Treeprasertsuk S, Tangkijvanich P, Payungporn S. Serum miR-29a and miR-122 as potential biomarkers for non-alcoholic fatty liver disease (NAFLD). *Microna* 2018;7:215–222.
- [36] Arrese M, Eguchi A, Feldstein AE. Circulating microRNAs: emerging biomarkers of liver disease. *Semin Liver Dis* 2015;35:43–54.
- [37] Roy S, Benz F, Vargas Cardenas D, Vucur M, Gautheron J, Schneider A, et al. miR-30c and miR-193 are a part of the TGF-beta-dependent regulatory network controlling extracellular matrix genes in liver fibrosis. *J Dig Dis* 2015;16:513–524.
- [38] Roy S, Hooiveld GJ, Seehawer M, Caruso S, Heinzmann F, Schneider AT, et al. microRNA 193a-5p regulates levels of nucleolar- and spindle-associated protein 1 to suppress hepatocarcinogenesis. *Gastroenterology* 2018;155:1951–1966.e1926.
- [39] Butt AM, Raja AJ, Siddique S, Khan JS, Shahid M, Tayyab GU, et al. Parallel expression profiling of hepatic and serum microRNA-122 associated with clinical features and treatment responses in chronic hepatitis C patients. *Sci Rep* 2016;6:21510.
- [40] Hyun J, Wang S, Kim J, Rao KM, Park SY, Chung I, et al. MicroRNA-378 limits activation of hepatic stellate cells and liver fibrosis by suppressing Gli3 expression. *Nat Commun* 2016;7:10993.
- [41] Zheng X, Zhao W, Ji P, Zhang K, Jin J, Feng M, et al. High expression of Rap2A is associated with poor prognosis of patients with hepatocellular carcinoma. *Int J Clin Exp Pathol* 2017;10:9607–9613.
- [42] Lopez MF, Zheng L, Miao J, Gali R, Gorski G, Hirschhorn JN. Disruption of the Igf2 gene alters hepatic lipid homeostasis and gene expression in the newborn mouse. *Am J Physiol Endocrinol Metab* 2018;315:E735–E744.
- [43] Hatzoglou M, Fernandez J, Yaman I, Closs E. Regulation of cationic amino acid transport: the story of the CAT-1 transporter. *Annu Rev Nutr* 2004;24:377–399.
- [44] Liu X, Xu J, Brenner DA, Kisseleva T. Reversibility of liver fibrosis and inactivation of fibrogenic myofibroblasts. *Curr Pathobiol Rep* 2013;1:209–214.
- [45] Sun J, Liu J, Zhu Q, Xu F, Kang L, Shi X. Hsa\_circ\_0001806 acts as a ceRNA to facilitate the stemness of colorectal cancer cells by increasing COL1A1. *Oncotargets Ther* 2020;13:6315–6327.
- [46] Povero D, Eguchi A, Li H, Johnson CD, Papouchado BG, Wree A, et al. Circulating extracellular vesicles with specific proteome and liver microRNAs are potential biomarkers for liver injury in experimental fatty liver disease. *PLoS One* 2014;9:e113651.
- [47] Ramming T, Hansen HG, Nagata K, Ellgaard L, Appenzeller-Herzog C. GPx8 peroxidase prevents leakage of H<sub>2</sub>O<sub>2</sub> from the endoplasmic reticulum. *Free Radic Biol Med* 2014;70:106–116.
- [48] Pagliassotti MJ. Endoplasmic reticulum stress in nonalcoholic fatty liver disease. *Annu Rev Nutr* 2012;32:17–33.
- [49] Khatib A, Solaimuthu B, Ben Yosef M, Abu Rmaileh A, Tanna M, Oren G, et al. The glutathione peroxidase 8 (GPX8)/IL-6/STAT3 axis is essential in maintaining an aggressive breast cancer phenotype. *Proc Natl Acad Sci U S A* 2020;117:21420–21431.
- [50] Liu CH, Ampuero J, Gil-Gómez A, Montero-Vallejo R, Rojas A, Muñoz-Hernandez R, et al. miRNAs in patients with non-alcoholic fatty liver disease: a systematic review and meta-analysis. *J Hepatol* 2018;69:1335–1348.