PNAS
www.pnas.org

Supplementary Information for

Identification of Pparγ-modulated miRNA hubs that target the fibrotic tumor microenvironment

Ivana Winkler, Catrin Bitter, Sebastian Winkler, Dieter Weichenhan, Abhishek Thavamani, Jan G. Hengstler, Erawan Borkham-Kamphorst, Oliver Kohlbacher, Christoph Plass, Robert Geffers, Ralf Weiskirchen, Alfred Nordheim

Alfred Nordheim
E-mail: alfred.nordheim@uni-tuebingen.de

This PDF file includes:

Figs. S1 to S13
Tables S1 to S4
Captions for Databases S1 to S5
References for SI reference citations

Other supplementary materials for this manuscript include the following:

Databases S1 to S5

Methods

Vectors pSV Sport PPAR gamma 1 and pMSCV-Lin28A were acquired from Addgene under the Uniform Biological Materials Transfer Agreement (UBMTA).

***SRF-VP16^{iHep}* animal model.** *SRF-VP16^{iHep}* mice carry and express a constitutively active form of SRF, SRF-VP16. The SRF-VP16 fusion protein is comprised of the first 412 residues of human SRF fused to the transcriptional activation domain of the *Herpes simplex* viral VP16 protein. The VP16 transcriptional activation domain enables SRF-directed target gene activation independently of upstream stimulation by Rho/actin and Ras/MAPK signaling pathways, thus rendering SRF-VP16 constitutively active (1).

To allow the conditional expression of SRF-VP16 upon cellular Cre recombinase activity, a floxed STOP cassette was introduced upstream of the SRF-VP16 coding sequence. The SRF-VP16 construct was integrated into the genomic Rosa26 locus generating *Gt(ROSA)26-Sor$^{tm1(SRF-VP16)Antu}$* mice.

To obtain *SRF-VP16^{iHep}* mice, stop-floxed *SRF-VP16* mice were bred with *Srf-flex1* mice (floxed *Srf* exon 1) and *Alfp-CreER^{T}2* animals which express tamoxifen-inducible hepatocyte-specific Cre recombinase. Therefore, this animal model shows conditional, Cre-mediated expression of SRF-VP16, combined with non-functional endogenous SRF. Tamoxifen treatment through activation of Cre recombinase efficiently induces SRF-VP16 expression, which leads to hepatocyte hyperproliferation (2).

Although Cre activation is dependent on tamoxifen treatment, spontaneous, stochastic Cre activation was also observed, leading to mosaic SRF-VP16 expression in a small number of hepatocytes. Hyperproliferation of affected hepatocytes leads to development of premalignant nodules throughout the liver and their gradual progression to HCC (2).

The tissue samples of premalignant nodules and HCC used in this study were isolated from *SRF-VP16^{iHep}* mice, which developed HCC as a result of stochastic hepatocyte-specific activation of SRF-VP16. Liver tissue from litter siblings that did not develop HCC due to lack of either SRF-VP16 or Cre expression was used as controls.

The murine carbon tetrachloride model. The murine CCl_4 model serves as an experimental *in vivo* liver fibrosis model. CCl_4 (Sigma-Aldrich) was administered intraperitoneally to C57BL/6 mice (n=5) twice per week for six weeks at a concentration of 0.8 μl/(g body weight) diluted in mineral oil following standardised operation procedures (3). Animals that received mineral oil alone served as controls.

Isolation of pHSC. Isolation of primary murine hepatic stellate cells (pHSC) from C57BL/6 mice was essentially performed as described previously (4). Briefly, healthy mice weighing about 20-25 g were sacrificed and the liver was perfused with pronase-collagenase solutions through the portal vein. The cells of the digested livers were dispersed and filtered through a nylon mesh. pHSCs were further enriched and purified from the remaining cells by Nycodenz (Axis-Shield) density centrifugation. Cells were counted in a Neubauer chamber and the viability of cells determined by trypan blue exclusion method using a ready-to-use 0.4% trypan blue solution. Cells were finally seeded on uncoated plastic in Dulbecco's modified Eagle's Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 100 IU/mL penicillin, 10 μg/mL streptomycin and 4 mM glutamine.

Activation of pHSC. pHSC activation occurs when cells are plated on standard tissue plastic plates (4). To ensure the full activation of pHSC, cells were maintained in DMEM culture medium supplemented with 10% FCS, 100 IU/mL penicillin, 10 μg/mL streptomycin and 4 mM glutamine for seven days. During this period cells were passaged once (after four days of cell culture) using the standard procedure for cell culture maintenance.

Cultivation of cell lines. The continuous murine hepatic stellate cell line GRX is an anchorage-dependent line displaying morphological characteristics of myofibroblasts (aHSC) (5). The cell line was obtained from the Rio de Janeiro Cell Bank (PABCAM, Federal University, Rio de Janeiro, Brazil). The murine fibroblast cell line NIH/3T3 (6), which when maintained under Tgfβ-rich conditions shows characteristics of myofibroblasts (7), was obtained from the American Type culture collection (ATCC). Both cell lines were maintained under standard growth conditions at 37°C in a humidified atmosphere with 5% CO_2. Routinely, cells were grown in DMEM culture medium supplemented with 10% FCS, 100 IU/mL penicillin, 10 μg/mL streptomycin and 4 mM glutamine.

Cell transfection. Cells were transfected using either electroporation or lipofection procedure.

To perform the electroporation, 10^6 cells were used per experiment. The electroporation pulse generator was set to an impulse of 320 V for 15 ms. To transfect the cells, 20 µg of vector DNA was used together with 100 µg of salmon sperm carrier DNA (Thermo Fisher Scientific). Following electroporation, cells were incubated for 24 h.

To transfect vector DNA, lipofection reagent TransIT-LT1 (Mirus Bio) reagent was used. To transfect miRNA mimics and miRNA inhibitors (Dharmacon) RNAiMAX (Thermo Fisher Scientific) and DharmaFECT1 (Dharmacon) reagents were used, respectively.

If miRNA mimics or inhibitors were to be transfected together with vector DNA (e.g. as in luciferase assays), the vector DNA was transfected first using the TransIT-LT1 Transfection protocol according to the manufacturer's instructions. At least 4 h post DNA transfection, the medium was changed and the RNAiMAX or DharmaFECT1 protocol was performed using miRNA mimics or miRNA inhibitors.

To generate stable cell lines, vector DNA was transfected using electroporation procedure. 24 h post-transfection, the cell culture medium was exchanged with a fresh medium supplemented with selection antibiotic (NIH/3T3 cells - 600 µg/ml G418 (Invivogen) and GRX cells - 700 µg/ml G418). Selection medium was changed every 2 to 3 days. Cells were expanded over 2 to 3 weeks to ensure the survival of only stably transfected cells.

Total RNA isolation. The mirVana isolation kit (Thermo Fisher Scientific) was used to prepare total RNA used for sRNA-seq. The isolation procedure was performed as stated in the manufacturer's instructions for isolation of total RNA. The RNeasy kit (Qiagen) was used to prepare RNA for RNA-seq according to the kit instructions. Alternatively, RNA isolation was performed using Trizol (Thermo Fisher Scientific) according to the manufacturer's instructions.

DNase treatment of RNA samples. To remove genomic DNA contamination in RNA samples, samples were treated with the RNAse-free DNase Treatment and Removal Reagents Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

DNA isolation. The DNA used in the methylation study, was isolated using DNeasy Blood & Tissue kit (Qiagen) according to the manufacturer's instructions.

sRNA-seq. To profile the whole miRNome, sRNA-seq was performed on tumor and nodular tissue alongside the corresponding controls. List of used samples and corresponding metadata are presented in Suppl. Table 1.

The library was generated from 1 µg of total RNA using TruSeq Small RNA Library Prep Kits v2 (Illumina) according to manufacturer's protocol. The libraries were sequenced on Illumina HiSeq2500 using TruSeq SBS Kit v3-HS (50 cycles, single ended run) with an average of 10^7 reads per RNA sample. FASTQ files were trimmed with Cutadapt (v1.11) removing Illumina RNA adapter sequences (TGGAATTCTCGGGTGCCAAGG) and nucleotides with PHRED scores below 20. For each FASTQ file a quality report was generated using FASTQC (v0.11.4) tool before and after trimming. Alignment of trimmed FASTQ sequences was done using STAR Aligner (v2.5.2b) (8) against the mouse genome GRCm38 using miRBase (v21) annotation. One mismatch was allowed for successful alignment in at least 15 matches per sequence. Counts per smallRNA (feature) were calculated using STAR's parameter *quantMode* set to*GeneCounts*. Differential expression of small RNAs was determined using DESeq2 (9) (v1.22.2).

The code for the bioinformatic analysis outlined here is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

RNA-seq. The samples used for RNA-seq analysis and corresponding metadata are listed in Suppl. Table 1.

The RNAseq library was generated using the TruSeq RNA sample preparation kit (Illumina) according to manufacturer's protocol.

The libraries were sequenced on Illumina HiSeq2000 producing 100 bp long, single ended reads. On average 10^7 reads were sequenced per RNA sample.

FASTQ files were trimmed with Cutadapt (v1.11) removing Illumina RNA adapter sequences (AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC) and nucleotides with PHRED scores below 20. For each FASTQ file a quality report was generated by FASTQC (v0.11.4) tool before and after trimming. Alignment of trimmed FASTQ sequences was done using STAR Aligner (v2.5.2b) against the mouse genome GRCm38. Counts per gene (feature) were calculated using STAR's parameter *quantMode* set to *GeneCounts*. Differential gene expression was determined using DESeq2 (v1.22.2).

The code for the bioinformatic analysis outlined here is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

Reverse transcription and quantitative PCR. Total RNA was reverse transcribed using miScript II RT Kit (Qiagen) which allows the conversion of all RNA species into cDNA.

The miScript SYBR Green PCR Kit (Qiagen) was used to quantify the expression of mature miRNAs and FastStart Universal SYBR Green Master (Roche) for pre-miRNA, pri-miRNA or mRNA quantification. To normalize the amount of cDNA between samples, endogenous controls were run alongside targets of interest. U6 small nuclear RNA (Rnu6) and/or small nucleolar RNA, C/D box 33 (Snord33) were used to normalize mature miRNA and the genes *Glyceraldehyde-3-phosphate dehydrogenase (Gapdh)*, *β-glucuronidase (Gusb)* and/or *TATA-box binding protein (Tbp)* served as endogenous controls for mRNA, pre-miRNA and pri-miRNA. Additionally, no-template controls were prepared for each primer pair as a control of cross-contamination. Each combination of cDNA and primer pair was analysed in triplicates in 10 µl reactions. All primer sequences are listed in supplementary material.

qPCR run was executed in QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher Scientific). To validate the specificity of PCR, the amplicon melting curve was determined for each reaction in every run.

Postrun data processing was conducted using LinRegPCR (10) (http://LinRegPCR.nl). Result of the LinRegPCR analysis are target gene's starting concentrations which can easily be compared between different samples. The relative expression of target genes/miRNAs was calculated by averaging N_0 values of technical replicates. To be able to compare the N_0 value of target genes/miRNAs across samples, the N_0 value of the targets was normalised to the N_0 value of the endogenous controls. The N_0 value of all samples are expressed relative to a randomly chosen control sample. Data processing and statistical analysis were performed using R (https://www.R-project.org, v3.5.1). Normalised N_0 values were compared using an two-sided unpaired t-test. Differences were considered statistically significant if the p-value was ≤ 0.05.

Staining of paraffin-embedded HCC samples. Formalin-fixed and paraffin-embedded tissue samples (tumors and the corresponding control tissue) were deparaffinized in Rotihistol (Carl Roth) and rehydrated in an ethanol gradient. The antigen retrieval step was performed using citrate buffer. For immunostaining, the MaxFlour 488 Immunofluorescence Detection Kit (Dianova) was used together with Early Growth Response (Egr1) (Cell Signaling, cat. 4153) and Alpha Smooth Muscle Actin (ACTA2) (Abcam, ab7817) antibodies. The Egr1 and ACTA2 antibodies were diluted 1:800 and 1:100, respectively. Nuclei were stained using DAPI (Thermo Fisher Scientific).

To visualise collagen depositions in the embedded tissue samples, Sirius Red staining was performed. Samples were de-waxed and hydrated. Nuclei were visualised by Weigert's haematoxylin. Subsequently, tissue slides were stained with Picro-Sirius Red (Sigma-Aldrich, Cat#365548) and washed in acidified water. Samples were dehydrated in absolute ethanol, cleared in xylene and mounted in a resinous medium. Images were acquired by an Olympus confocal laser scanning microscope and quantified using ImageJ. The macro used for quantification is deposited in the GitHub repository (see *Data and code availability*). Measured stained %Area of total area was compared between control, nodular and tumor samples using one-way ANOVA and a post-hoc Tukey test.

Chromatin immunoprecipitation. The used ChIP protocol is based on the procedure described by Daniel et al. (11) with some modifications.

To reversibly cross-link the adherent cell lines, formaldehyde (Thermo Fisher Scientific) was directly added on the cell plate in a final concentration of 1 %. Per preparation (sample) 30 million cells were used. Added formaldehyde was diluted in phosphate-buffered saline (PBS). Cells were incubated for 10 min and the cross-linking reaction was stopped by adding glycine (Applichem) to a final concentration of 0.125 M.

Cells were scraped off the plate in 1 ml of cell lysis/wash buffer (0.15 M NaCl, 0.005 M EDTA pH 7.5, 0.05 M Tris–HCl pH 7.5, 0.5 % NP40, dH_2O supplemented with protease inhibitor (Roche) prior to use) and pelleted at 12,000 x g for 1 min. The pellet was resuspended in additional 1 ml of the cell lysis/wash buffer and passed through an insulin syringe. Following the centrifugation step, nuclei pellets were resuspended in 700 µl of nuclear lysis buffer (50 mM Tris-HCl pH 7.5, 1 % SDS, 20 mM EDTA).

Chromatin, which was released in the previous step, was subsequently fragmented by sonication into fragments of 200-1000 base pairs (bp). Sonication was performed for 15 pulses of 20 s with a 30 s resting interval at 4°C on high energy settings (Bioruptor, Diagenode). Efficiency of sonication was validated by loading a small aliquot of chromatin onto the agarose gel. Remaining chromatin was cleared by centrifugation at 12,000 x g at 4°C for 10 min.

10% of chromatin was used as input. Chromatin used in immunoprecipitation was diluted 10 times using ChIP Dilution Buffer (0.001 M EDTA, pH 8.0, 0.017 M Tris–HCl, pH 8.0, 0.01 % SDS, 1.1 % Triton-X 100, 0.17 M NaCl, dH_2O supplemented with protease inhibitor tablets (Roche) prior to use).

Chromatin was split in two fractions. The first fraction was incubated with 2 µg antibody of interest (PPARγ (Perseus proteomics, PP-A3409A)) and the second with IgG antibody (Milipore) used to assess the unspecific antibody binding. Samples were incubated overnight at 4 °C.

The following day, Protein A-coupled Dynabeads (Thermo Fisher Scientific) were blocked using Blocking buffer (0.1% PVP, 0.1% UltraPure BSA in PBST (0.1% TWEEN-20 in PBS)) for 30 min at room temperature.

The blocked bead suspension was added to the chromatin sample and chromatin-antibody-bead complexes were incubated for 6 h at 4 °C.

The samples were washed using cell lysis/wash buffer five times. Chromatin was eluted in 450 µl Elution buffer at 65°C. To reverse the cross-linking, 20 µl 5M NaCl and 20 µl 0.5 M EDTA were added to all samples and samples were incubated at 65 °C overnight.

DNA was isolated using a standard phenol:chloroform:isoamyl alcohol procedure. Target protein-bound DNA in immunoprecipitated samples, together with input samples, was quantified in qPCR. The Percent Input Method was used as method of normalization. Calculated values of Percent Input were compared using an two-sided unpaired t-test. Differences were considered statistically significant if the p-value was ≤ 0.05.

Methylation analysis. Quantitative DNA methylation analysis was performed by matrix-assisted time-of-flight mass spectrometry (MassARRAY; Agena Bioscience) essentially as described previously (12) using primers listed in supplementary material. In brief, genomic DNA was bisulfite-converted and used as template to generate PCR amplicons with a T7-promoter tag. *In vitro*-generated RNA from the amplicons was cleaved by RNase A providing specific fragments which were analysed by MassARRAY. Fragments with ambiguous or too low ($\leq$1500) or too high ($\geq$7000) masses were omitted from the analysis. Beta values were compared using Mann–Whitney U test. Differences were considered statistically significant if the p-value was ≤ 0.05.

Luciferase assay. To experimentally validate the functionality of predicted miRNA targeting, a luciferase gene reporter assay was used. The predicted target site or the full length 3'-UTR and a mutated version were subcloned downstream of the luciferase gene and transfected together with a specific miRNA mimic or inhibitor. The list of miRNA:mRNA pairs for which targeting was experimentally validated and obtained corresponding experimental data are presented in Suppl. Table 2.

pmirGLO Dual-Luciferase miRNA Target Expression Vector was used as backbone vector to clone in a partial or full length 3'-UTR region of *Col1a1, Pdgfa, Adamts15* and *Tgfbr1* genes. Following vector transfection, cells were transfected with a miRNA mimic or inhibitor of interest. Final concentration of all miRNA mimics and inhibitors was 50 µM.

To mutate miRNA sites in all pmirGLO vectors, Q5 Site-Directed Mutagenesis kit (NEB) was used according to the manufacturer's instructions.

After 24 h, luciferase activity was measured using Dual-Luciferase Reporter Assay (Promega) according to the manufacturer's instructions. Luciferase activity was measured using OPTIMA FluoroSTAR. To compare measured luciferase flourescence units (LFU), a two-sided unpaired t-test was used. Differences were considered statistically significant if the p-value was ≤ 0.05.

Primers used for cloning and mutagenesis are listed in Additional data table S4.

Modulation of target gene expression using miRNA mimics and inhibitors. Monitoring of the target gene expression upon miRNA mimic or inhibitor transfection was used as alternative approach to experimentally validate the functionality of predicted miRNA targeting. miRNA mimics and inhibitors were transfected as described in section *Cell transfection*. Final concentrations of miR-29c and let-7g mimics were 100 µM, while miR-338 and let-7c mimics as well as let-7g and miR-29c inhibitors were used in final concentration of 50 µM. Gene expression was quantified as reported in section *Reverse transcription and quantitative PCR*.

PGJ_2 treatment of stable Pparγ-overexpressing GRX cells. To assess effects of Pparγ-mediated miRNA expression on fibrotic target genes, stable Pparγ-overexpressing GRX cells were treated with the Pparγ-agonist PGJ_2 (Sigma Aldrich) and simultaneously transfected with miRNA inhibitors. Transfection of miR-29c and let-7g inhibitors was performed as described in section *Cell transfection* using starvation medium (0.5% FCS, 100 IU/mL penicillin, 10 μg/mL streptomycin and 4 mM glutamine) with or without PGJ_2 (final concentration 2µM) as transfection medium. Cells were washed with PBS before transfection medium was added. Cells were incubated for 16 h before gene expression was quantified as reported in section *Reverse transcription and quantitative PCR*.

Vectors. To generate the pMSCV vector, the pMSCV-Lin28A vector (Addgene) was digested using *EcoRI* (Thermo Fisher Scientific) and *BglII* (Thermo Fisher Scientific) restriction enzymes under standard conditions. Digested vector DNA was loaded onto 1 % agarose gel and the vector backbone was isolated from the gel using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions.

Two single-stranded, complementary oligonucleotides were annealed to form a double-stranded DNA fragment. 3 µg of complementary oligonucleotides were mixed in 60 µl annealing buffer (100 mM NaCl, 50 mM HEPES, pH 7.4). Oligonucleotides were initially heated to 90°C for 4 min to dissolve secondary structures and gradually cooled to 37°C (0.03°C/s) to facilitate hybridization.

Annealed nucleotides were inserted into the vector backbone following the standard ligation protocol.

pSV Sport PPAR gamma 1 vector was modified by inserting the *Neomycine resistance gene* into the vector backbone.

pMSCV-mLin28A was a gift of George Daley (Addgene plasmid #26357) (13) and pSV Sport PPAR gamma 1 was kindly provided by Bruce Spiegelman (Addgene plasmid #8886) (14).

pmirGLO vector was purchased from Promega.

Short description of all vectors used in this study is contained in supplementary material (Suppl. Table 3).

miRNA target prediction. To identify miRNA targets, the DIANA microT-CDS (v5) (15) and TargetScan (v7.2) (16) databases were used. miRNA:mRNA targeting pairs, identified using aforementioned algorithms, were further filtered in accordance with their expression in our sRNA-seq and RNA-seq mouse datasets. miRNAs typically downregulate their targets. Thus, in order to identify targets of downregulated miRNAs, target mRNA candidates generated through bioinformatic analysis were matched with the upregulated genes found by RNA-seq. Similarly, upregulated miRNAs were matched with downregulated target mRNA candidates found by RNA-seq. miRNAs were considered down/upregulated when their expression was 1.5 fold down/upregulated with p_{adj}-value ≤ 0.05 in tumor compared to control. mRNAs were considered down/upregulated when their expression was two fold down/upregulated and/or p_{adj}-value ≤ 0.05. To profile evolutionary conserved targeting, a similar analysis was performed for the human HCC TCGA dataset. miRNAs and mRNAs in the TCGA dataset were considered to be dysregulated if at least one miRNA family member was dysregulated (threshold 1.5 fold) in $\geq 10\%$ of cases of the cohort. miRNA:target mRNA pairs of both datasets were overlapped and only those pairs that showed conservation (i.e. targeting in both mice and humans) were used in the gene enrichment analysis.

Gene enrichment analysis using predicted miRNA gene targets was performed using KEGG pathways (v6.2) and Reactome (v67). To perform gene enrichment analysis, over-representation analysis using a hypergeometric test was performed.

Transcription start site prediction of miRNA-encoding genes. A bioinformatic approach was used to identify TSSs of miRNA-encoding genes of interest. GRO-seq shows sharp peaks around TSSs in both the sense and antisense directions and a continuous signal of lower intensity throughout the entire transcript allowing to map TSS of transient transcripts. To map miRNA TSSs, six human (SRR014283, SRR574824, SRR1015583, SRR1145822, SRR1745515, SRR2961002) (17–21) and nine mouse (SRR097858, SRR097863, SRR097864, SRR1517780, SRR1772450, SRR1991266, SRR3051599, SRR3051601, SRR5816144) (22–27) GRO-seq datasets deposited in the Gene Expression Omnibus (GEO) were used. To download the datasets the FASTQdump (v2.8.2) tool was used, FASTQC (v0.11.4) tool to examine the quality of the datasets, STAR (v2.5.2b) to perform the read alignment and HOMER (v4.10) to perform peak analysis (28). Integration of miRNA TSSs from all datasets resulted in the list of unique miRNA TSSs.

Transcription factor binding site prediction. As promoter regions of miRNA-encoding genes, we considered regions 1000 bp downstream of a TSS and 500 bp upstream of a TSS. To identify transcription factors which can potentially bind to the miRNA-encoding gene promoters, FIMO (v4.12.0), part of the MEME suite, was used (29). JASPAR CORE (v7), which contains a curated, non-redundant set of profiles derived from published and experimentally defined transcription factor binding sites for eukaryotes, was used as motif database (30). To refine the prediction two RNA-seq datasets (GSE78853 - inactive (vehicle treated) and activated (TGF-β treated) hepatic stellate cells (31) and GSE93313 - quiescent and inactive pancreatic stellate cell lines) (32) were used. Dataset GSE78853 was downloaded as raw reads and processed using DESeq2 (v1.22.2), while dataset GSE93313 was downloaded as table of differentially expressed genes. The code for the bioinformatic analysis outlined here is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

TCGA data processing. The data used for this analysis were generated by the TCGA Research Network: http://cancergenome.nih.gov/.

TCGA's sRNA-seq, RNA-seq, methylation data and copy number variation (CNV) data for HCC, LUAD, LUSC and BRCA together with the corresponding metadata were downloaded from the TCGA data portal.

sRNA-seq and RNA-seq data were downloaded as raw reads and processed further using DESeq2 (v1.22.2). As not every tumor sample has a corresponding control sample, normalised reads of control samples were averaged and reads of individual tumor samples were compared to this average control value. To generate normalised reads, DESeq2 package was used.

CNV probe values were mapped to genes and normalised to the value of 2, which was taken as reference gene copy number. To process CNV data, R package CNTools was used.

Beta values of probes in the methylation dataset were mapped to the genes. Beta values of control samples were averaged and compared to beta values of individual tumor samples.

The code for the bioinformatic analysis outlined here is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

Analysis of miRNA-mRNA interactions across fibrotic cancers. To assess if a subset of miRNAs regulates common mRNAs across diverse cancer types, the multivariate linear regression approach introduced by *Jacobsen et al.* (33) was used.

For every pair (k, l) of mRNA k and miRNA l of interest and a given tumor type, the following regression model was employed. Let $y = (y_i) \in \mathbf{R}^n, i = 1, ..., n$ be the expression of mRNA k for all n samples of a given cancer type. y is then modeled as the dependent variable in a linear regression with CNV intensities, $x_{cnv} = (x_{cnv,i})$, DNA methylation beta fold changes, $x_{me} = (x_{me,i})$ and expression of miRNA l, $x_l = (x_{l,i})$ as independent variables where $i = 1, ..., n$ again denote the n tumor samples of a given cancer type:

$$y_i = \beta_0 + \beta_{cn} x_{cn,i} + \beta_{me} x_{me,i} + \beta_l x_{l,i} \quad , \quad i = 1, .., n \qquad [1]$$

Here, β_0 is the intercept, β_{cn}, β_{me} and β_l are the regression coefficients for the CNV, methylation and miRNA expression covariates, respectively. The model was fitted using the standard assumptions of multivariate regression with the Python (v3.6.7) package *statsmodels* (v0.9.0). As described in (33) a t-test can be used to asses miRNA regulation influence and a rejection of the null hypothesis $H_0 : \beta_l = 0$ at $\alpha = 0.05$ after Benjamini-Hochberg multiple testing correction indicates a regulation of mRNA k by miRNA l.

The rank-statistic approach described in (33) was used to evaluate the relative strength of miRNA:mRNA association across the different cancer types, giving an indication of consistent dysregulation patterns of miRNA:mRNA pairs.

For a given cancer type c, mRNA k and miRNA l, let $L_{l,c} = (r_{k,l,c})_k$ be the ordered list of miRNA regression coefficients across all mRNAs, from largest negative to largest positive. Then, (33) defines the rank $rr_{k,l,c}$ of mRNA-miRNA pair (k, l) in cancer type c as

$$rr_{k,l,c} = \frac{r_{k,l,c}}{|L_{l,c}|} - \frac{1}{2 \cdot |L_{l,c}|} \qquad [2]$$

Under the assumptions detailed by (33) the one-sided null hypothesis that no negative association of mRNA-expression and miRNA-expression exists can be formulated as $H_0 : -2 \sum_c ln(rr_{k,l,c}) \sim \chi^2_{2n}$ where χ^2_{2n} denotes a chi-squared distribution with $2n$ degrees of freedom. For every pair mRNA-miRNA pair (k, l) we thus obtain a p-value $p_+^{k,l}$ representing the strength of negative association of expression. Similarly, by reversing the ranks $rr_{k,l,c}$ via $|L_{l,c}| - rr_{k,l,c} + 1$ one obtains a p-value $p_-^{k,l}$ measuring positive association of expression across cancer types. $p_-^{k,l}$ and $p_+^{k,l}$ for all mRNA-miRNA pairs again were corrected for multiple hypothesis testing via Benjamini-Hochberg. (33) then define the REC-score (cross-cancer association **rec**urrence score) for every mRNA-miRNA pair as follows:

$$REC_{k,l} = \begin{cases} 2\, log_{10}(p_-^{k,l}) & \text{if } p_-^{k,l} < p_+^{k,l} \\ -2\, log_{10}(p_+^{k,l}) & \text{if } p_+^{k,l} < p_-^{k,l} \\ 0 \quad otherwise \end{cases} \qquad [3]$$

Correspondingly, a negative REC-score represents an expected miRNA:mRNA expression association pattern, i.e. upregulated miRNA and downregulated mRNA or downregulated miRNA and upregulated mRNA expression.

Processing of the TCGA data used to fit the model is described in section *TCGA data processing*.

The code for the bioinformatic analysis outlined here is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

Data analysis and visualisation. To perform statistical analysis, R (https://cran.r-project.org/) was used. R packages, used for specific bioinformatic analysis, are listed in the respective section where the used methods are described. Majority of the graphs were plotted using the ggplot2 package, with exception of the circos plot, for which the circlize package was used. To generate the model sub-figure (Figure 8), the Servier medical art collection (https://smart.servier.com/) was used.

Statistical analyses of data acquired are described in the Method detail section of each experimental method. A summary of statistical analyses is shown in Additional data table S1.

Ethics approval and consent to participate. Experiments involving *SRF-VP16*^{iHep}mouse model were approved by Regierungspräsidium Tübingen (IM1/14 permit). Experiments involving CCl_4 animal model complied with the guidelines for animal care and were approved by the German Animal Care Committee and the Landesamt für Umwelt und Naturschutz (LANUV, Recklinghausen, Germany) under permit no. Az.: 84-02.04.2012.A092. Permission to isolate pHSC from mice using the perfusion protocol was given by the LANUV under permit no.: Az.: 84.02.04.2015.A028.

Availability of data and materials. The code for the bioinformatic analysis outlined in the Material and method part is available in the following url: https://ivanawinkler.github.io/mirna_paper/.

sRNA-seq and RNA-seq FASTQ data are deposited in NCBI Sequence Read Archive (SRA) under following accession number, Bioproject: PRJNA522967. Processed sRNA-seq and RNA-seq (differential gene and miRNA expression table and normalised reads) are listed in supplementary material. Output data of the linear regression analysis are, as well, listed in Supplementary material.

Stable Pparγ-overexpressing GRX cell lines and stable Lin28a-overexpressing NIH/3T3 are available upon request. Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding author Alfred Nordheim (alfred.nordheim@uni-tuebingen.de).

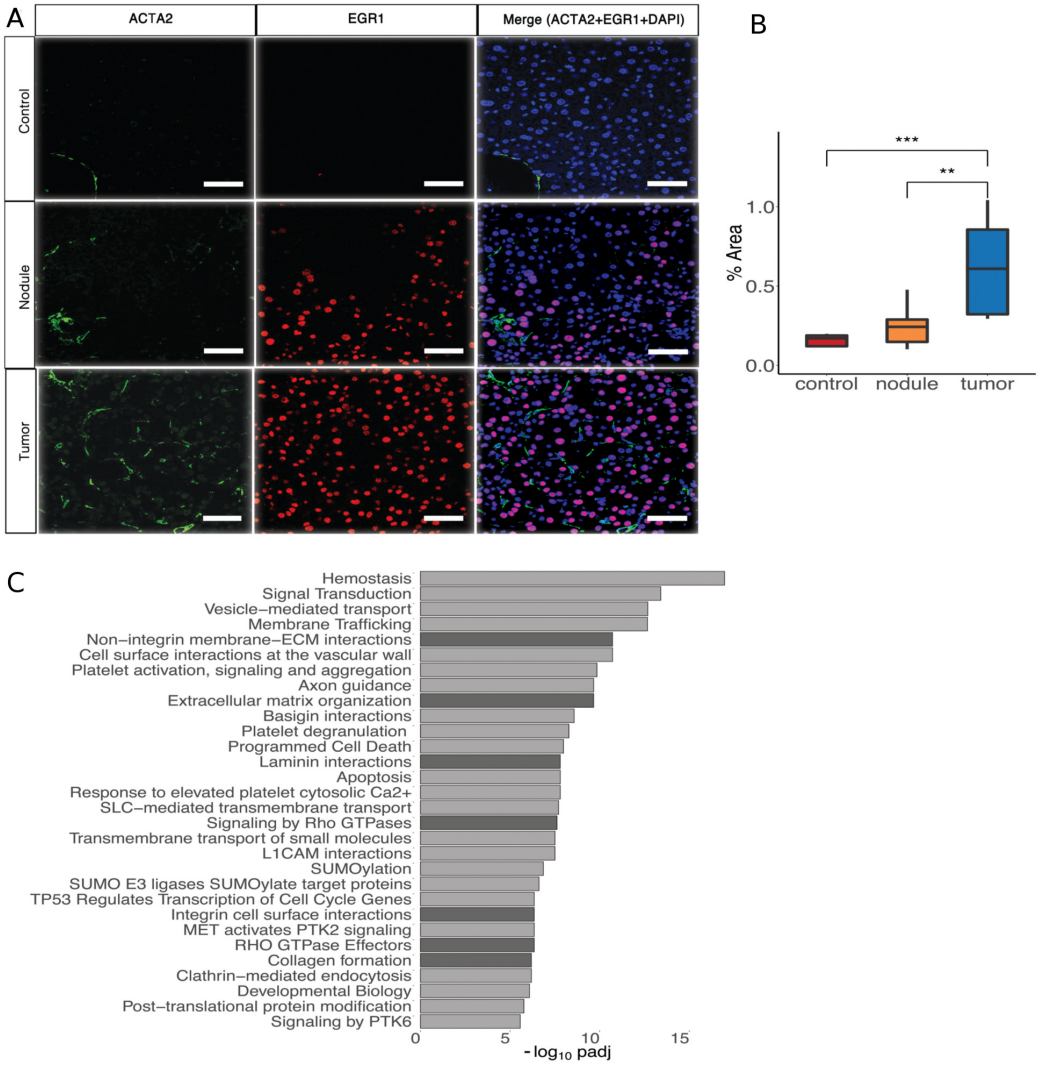

Fig. S1. A subset of miRNAs targets ECM-linked and fibrosis-associated genes in mHCC. (A) Acta2 (aHSCs), Egr1 (SRF-VP16 expressing hepatocytes) and merge (all cells) of control, nodular and tumor liver samples isolated from *SRF-VP16*^{iHep} mice. Scale 50 μm. (B) Quantification of Acta2 signal shown in B. (C) Gene set enrichment analysis of target genes of downregulated miRNAs found to be conserved in murine (*SRF-VP16*^{iHep}) and human (TCGA) HCCs. Pathways depicted by bars in dark grey represent ECM- and Rho GTPase-related pathways. Data are shown as median, first and third quartile ("box") and 95% confidence interval of median ("whiskers"). * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

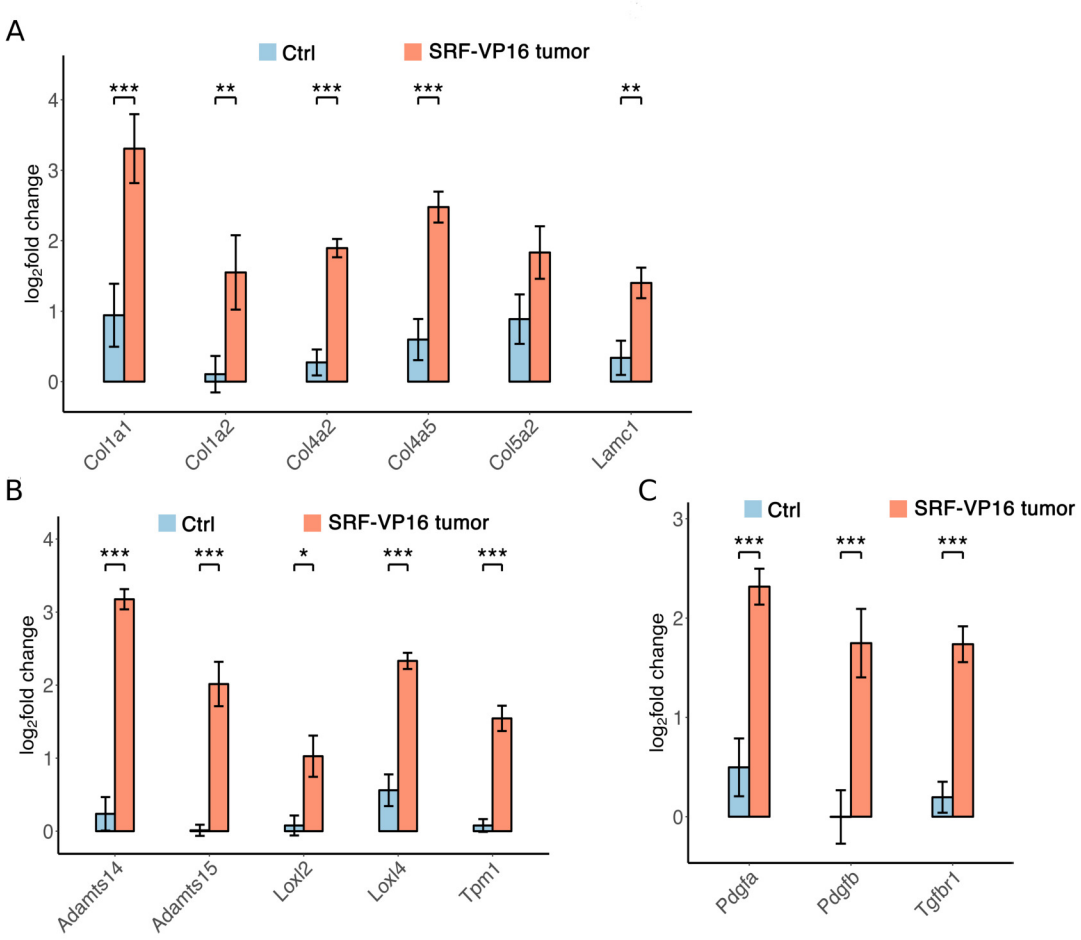

Fig. S2. AF-miRNAs are downregulated and fibrosis-associated genes are upregulated in murine HCC. (A-C) Normalized read count ($\log_2$-transformed) of fibrosis-associated, structural (A), remodeling (B) and signaling (C) genes of the ECM in control and tumor samples of *SRF-VP16*^{iHep} mice. All samples are normalized to a randomly chosen control sample. Data are shown as mean and standard error of the mean. p_{adj}-value $\leq$ 0.05, ** p_{adj}-value $\leq$ 0.01, *** p_{adj}-value $\leq$ 0.001.

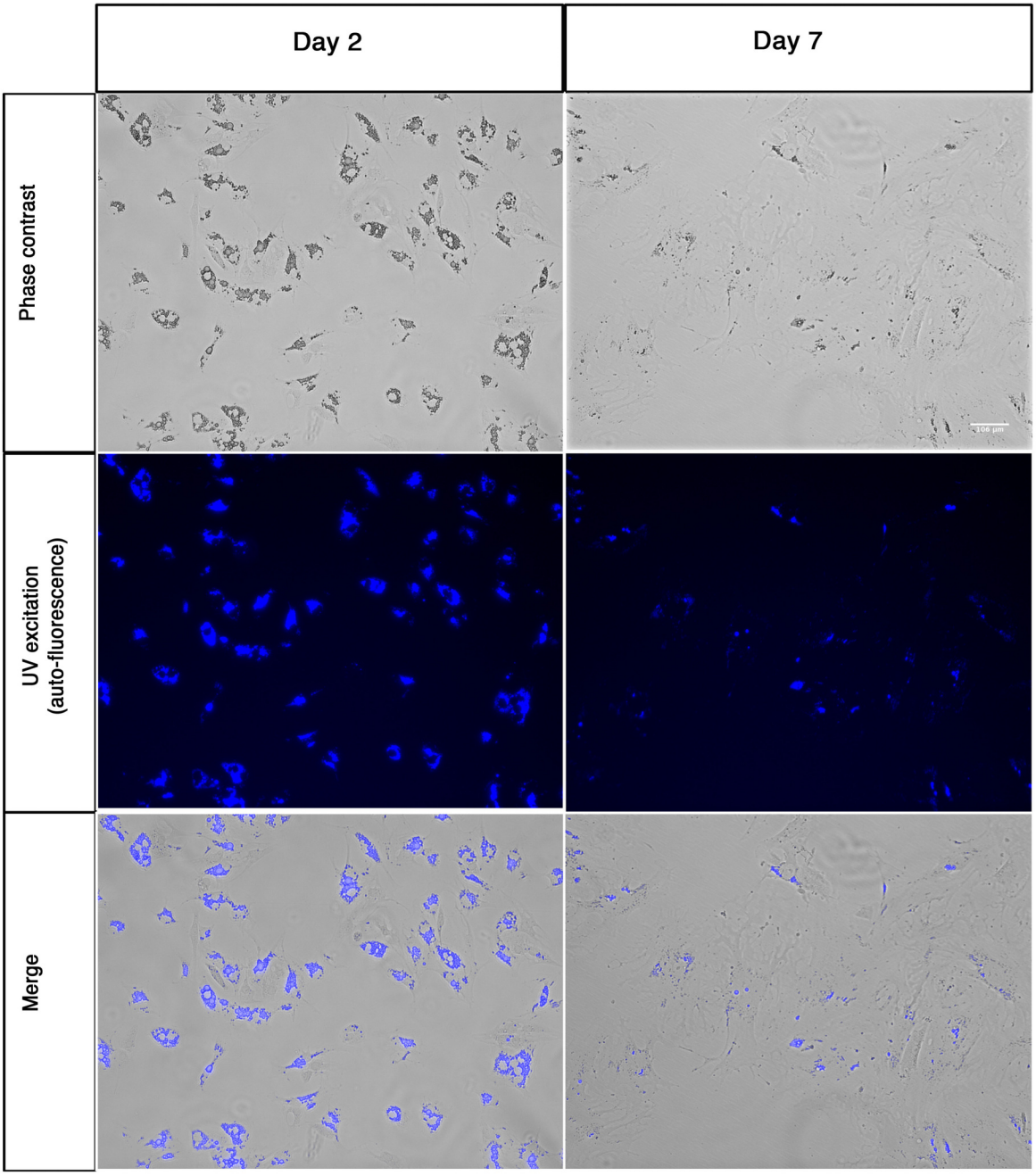

Fig. S3. pHSC are spontaneous activated as a result of prolonged growth in standard plastic cell culture dishes. Upper row shows changes in morphology of pHSCs upon culturing, as visualized by phase contrast microscopy. Middle row shows auto-fluorescence of retinoid droplets in cytoplasm of pHSCs upon UV excitation. Auto-fluorescence of retinoid droplets is diminished at day 7 of pHSC culturing, indicating activation of pHSCs. Bottom row shows merge of upper and middle rows.

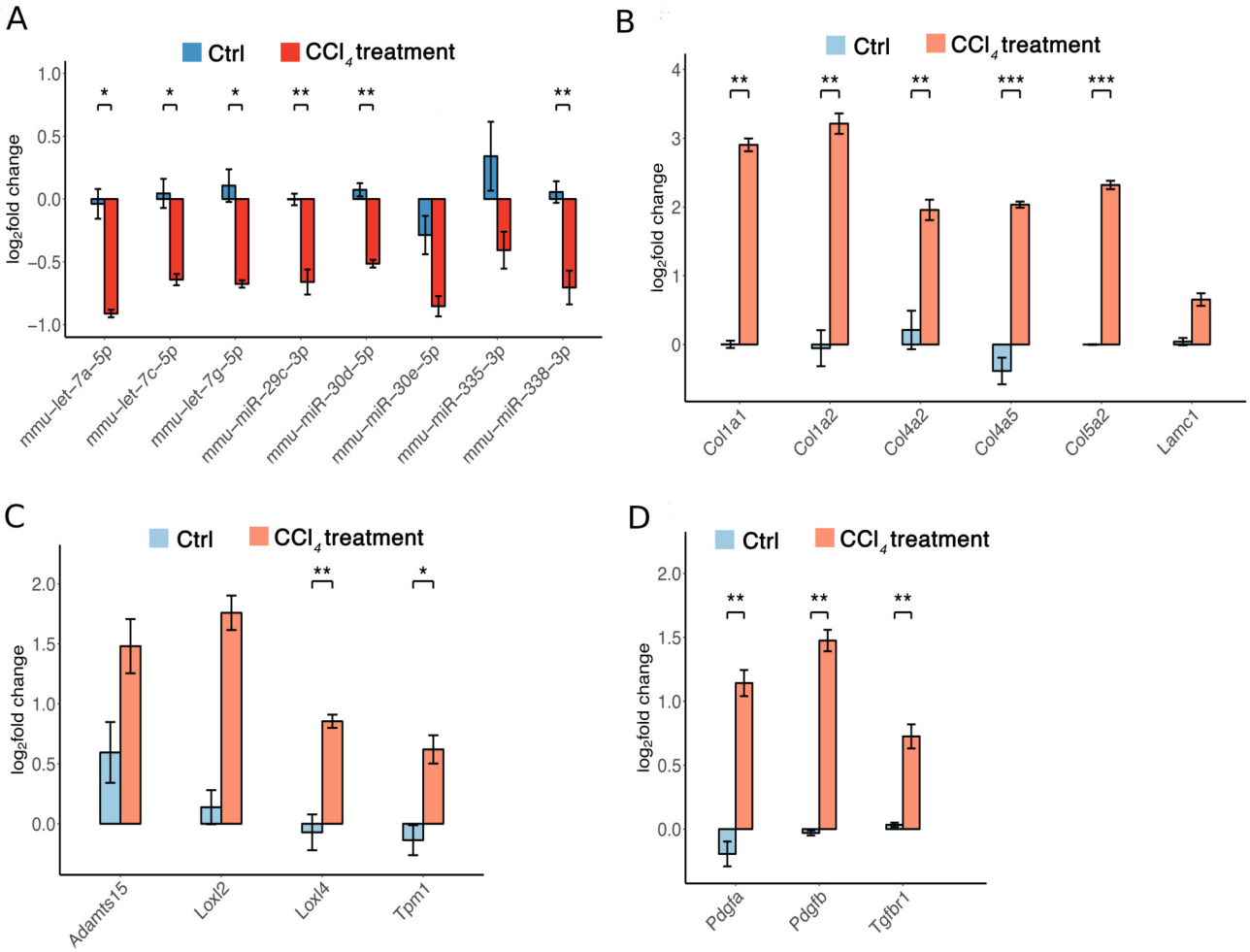

Fig. S4. AF-miRNAs are downregulated and fibrosis-associated genes are upregulated in mouse CCl_4 fibrosis models. (A) Relative expression of mature miRNAs in samples of CCl_4-treated mice in comparison to controls (mineral oil treatment). (B-D) Relative expression of fibrosis-associated, structural (B), remodeling (C) and signaling (D) genes of the ECM in CCl_4-treated and control mice. All samples are normalized to a randomly chosen control sample. Data are shown as mean and standard error of the mean. * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

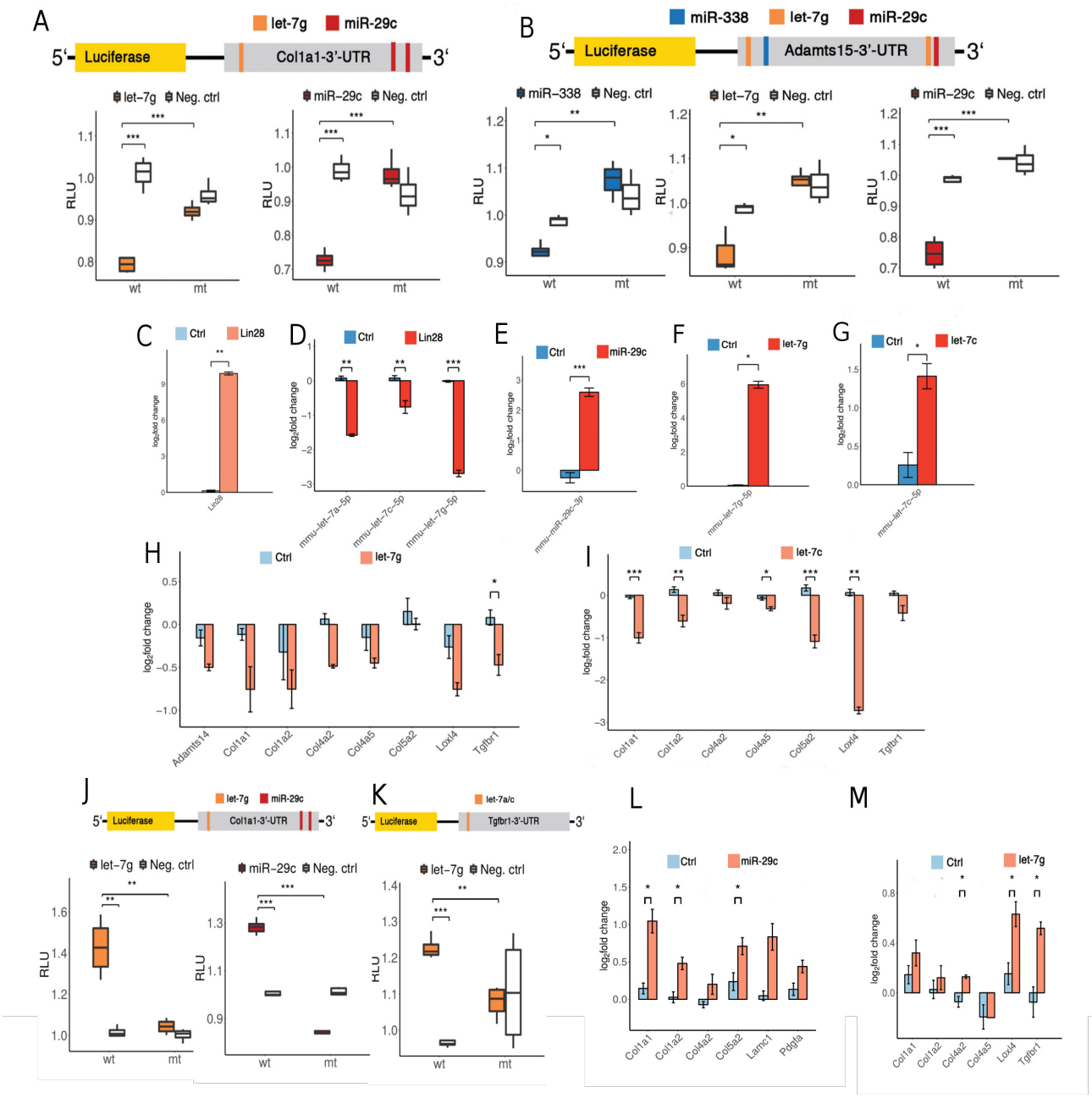

Fig. S5. AF-miRNAs target structural, signaling and remodeling components of the ECM. (A-B) Activities of wild-type and mutant (mutated miRNA site) 3'-UTR-luciferase constructs derived from: (A) *Col1a1* in NIH/3T3 cells transfected with let-7g, miR-29c and scrambled miRNA mimic (Neg. ctrl) and (B) *Adamts15* in NIH/3T3 cells transfected with miR-338, let-7g, miR-29c and scrambled miRNA mimic. Let-7g-, miR-29c- and miR-338-transfected samples are colored in the plots according to the luciferase construct schematic. Samples transfected with scrambled miRNA mimic are shown in white. (C-D) Relative expression of: (C) *Lin28a*, (D) let-7a, let-7c and let-7g in stable Lin28a-overexpressing NIH/3T3 cells. (E-G) Relative expression of (E) miR-29c, (F) let-7g and (G) let-7c in NIH/3T3 cells transfected with miR-29c, let-7g or let-7c mimics, respectively. (H-I) Relative expression of putative (H) let-7g and (I) let-7c target genes associated with fibrosis in NIH/3T3 cells transfected with let-7g and let-7c mimics, respectively. (J-K) Activities of wild-type and mutant (mutated miRNA site) 3'-UTR-luciferase constructs derived from: (J) *Col1a1* in NIH/3T3 cells transfected with miR-29c inhibitors, let-7g inhibitors and scrambled miRNA inhibitors (Neg. ctrl) and (K) *Tgfbr1* in NIH/3T3 cells transfected with let-7g inhibitors and scrambled miRNA inhibitors. (L-M) Relative expression of putative (L) miR-29c and (M) let-7g target genes associated with fibrosis in NIH/3T3 cells transfected with inhibitors of miR-29c and let-7g, respectively. (A-B, J-K) Data are shown as median, first and third quartile ("box") and 95% confidence interval of median ("whiskers"). (C-I, L-M) Data are shown as mean and standard error of the mean. * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

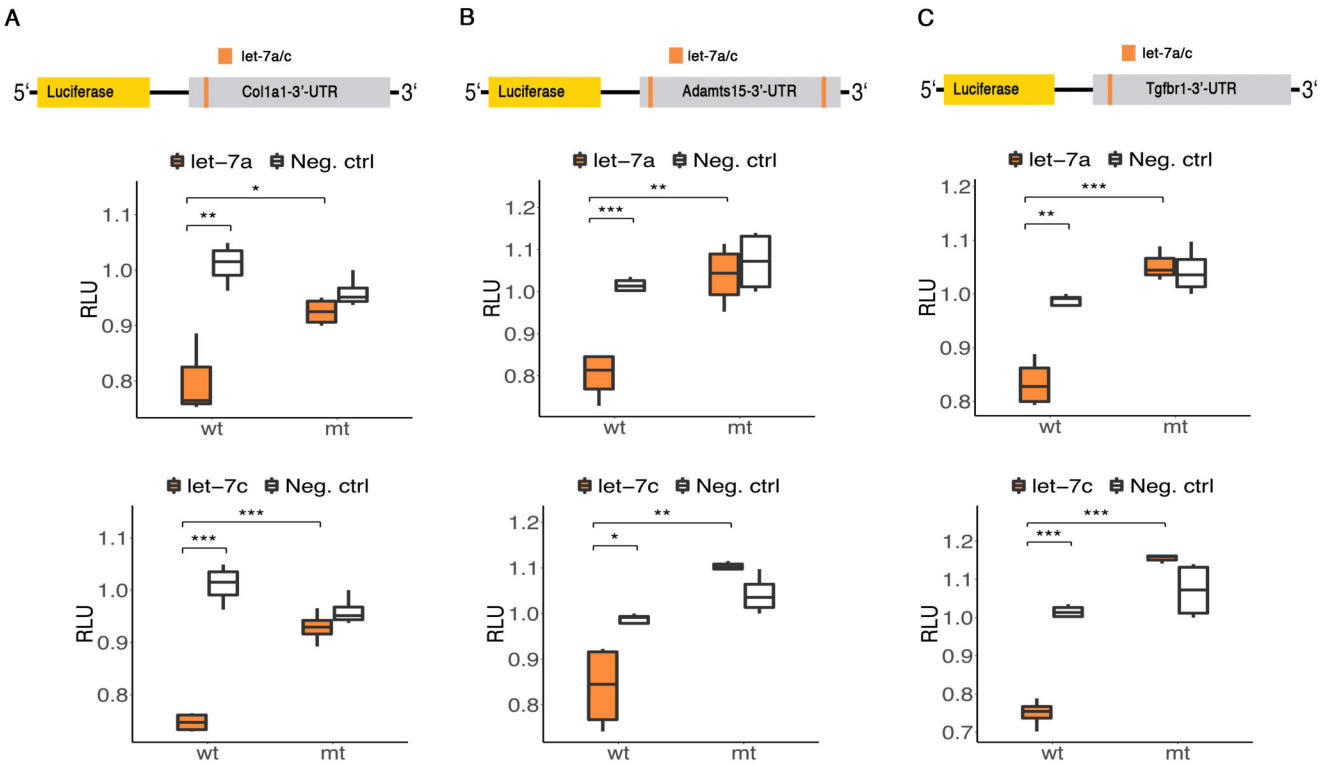

Fig. S6. Let-7 targets fibrosis-associated genes. (A-C) Measurement of the activities of wild-type and mutant (mutated miRNA interaction site) 3'-UTRs luciferase constracts derived from: (A) *Col1a1*, (B)*Adamts15* and (C) *Tgfbr1* genes in NIH/3T3 cells transfected with let-7a, let-7c and scrambled miRNA mimic (Neg. ctrl). Data are shown as median, first and third quartile ("box") and 95% confidence interval of median ("whiskers"). * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

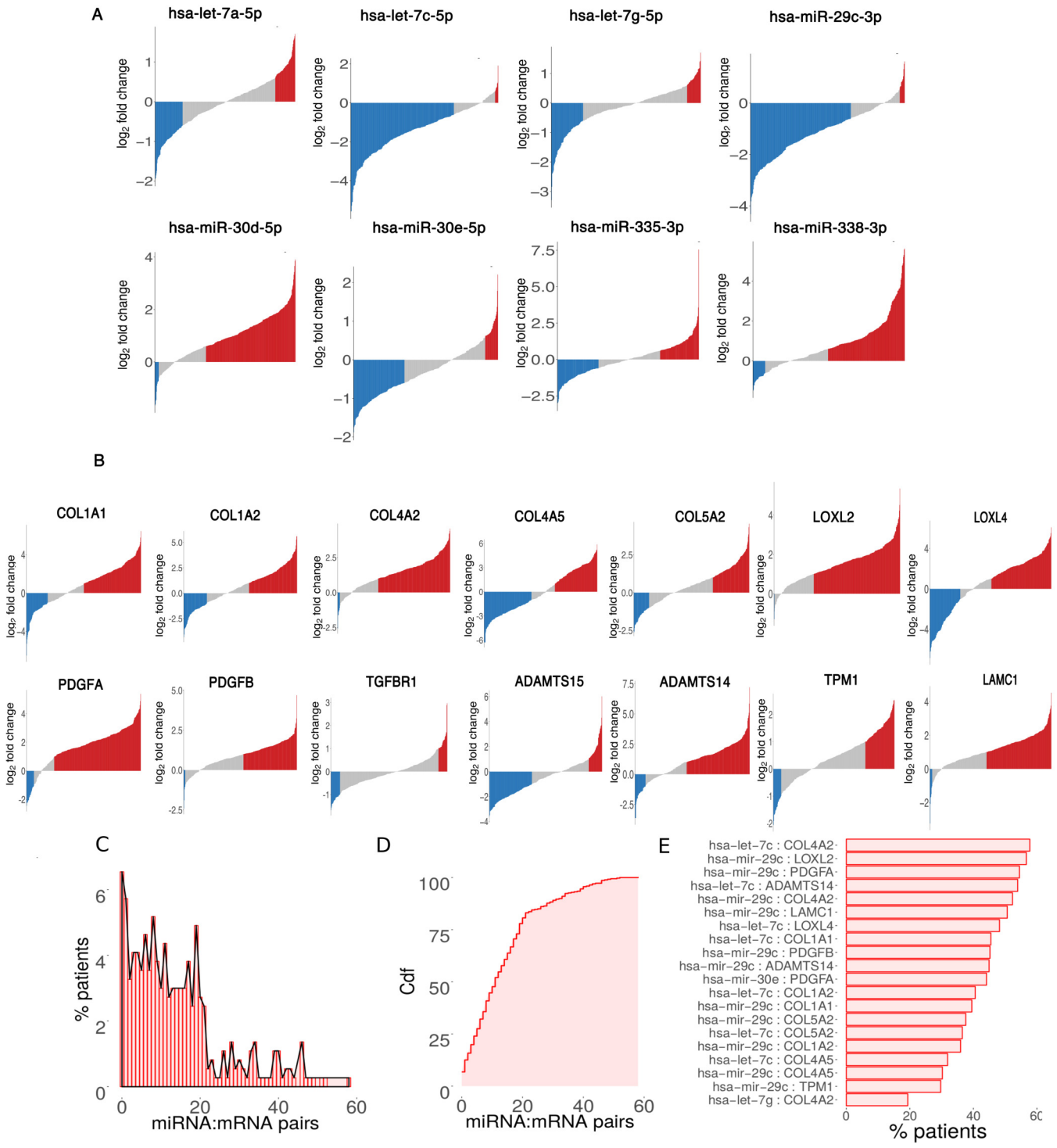

Fig. S7. AF-miRNAs are downregulated and fibrosis-associated genes are upregulated in a subset of human HCCs. (A) Log_2-fold changes of AF-miRNAs in individual tumors (compared to the mean control value) of the human TCGA dataset. Bars in waterfall plot: blue (miRNA downregulation $\geq$ 1.5-fold), red (miRNA upregulation $\geq$ 1.5-fold). (B) Log_2-fold changes of fibrosis-associated genes in individual tumors (compared to the mean control value) of the human TCGA dataset. Bars in waterfall plot: blue (mRNA downregulation $\geq$ 2-fold), red (mRNA upregulation $\geq$ 2-fold). (C) Number of AF-miRNA and fibrosis-associated gene (mRNA) pairs conserved in each individual patient. (D) Cumulative distribution function (Cdf) of patients with a number of conserved AF-miRNA and fibrosis-associated gene (mRNA) pairs. (E) Top 20 AF-miRNA and fibrosis-associated gene (mRNA) pairs conserved across patients of the TCGA cohort. (C-E) miRNAs are considered downregulated if their expression is $\geq$ 1.5-fold downregulated in comparison to mean control value and mRNAs upregulated if their expression is $\geq$ 1.5-fold upregulated.

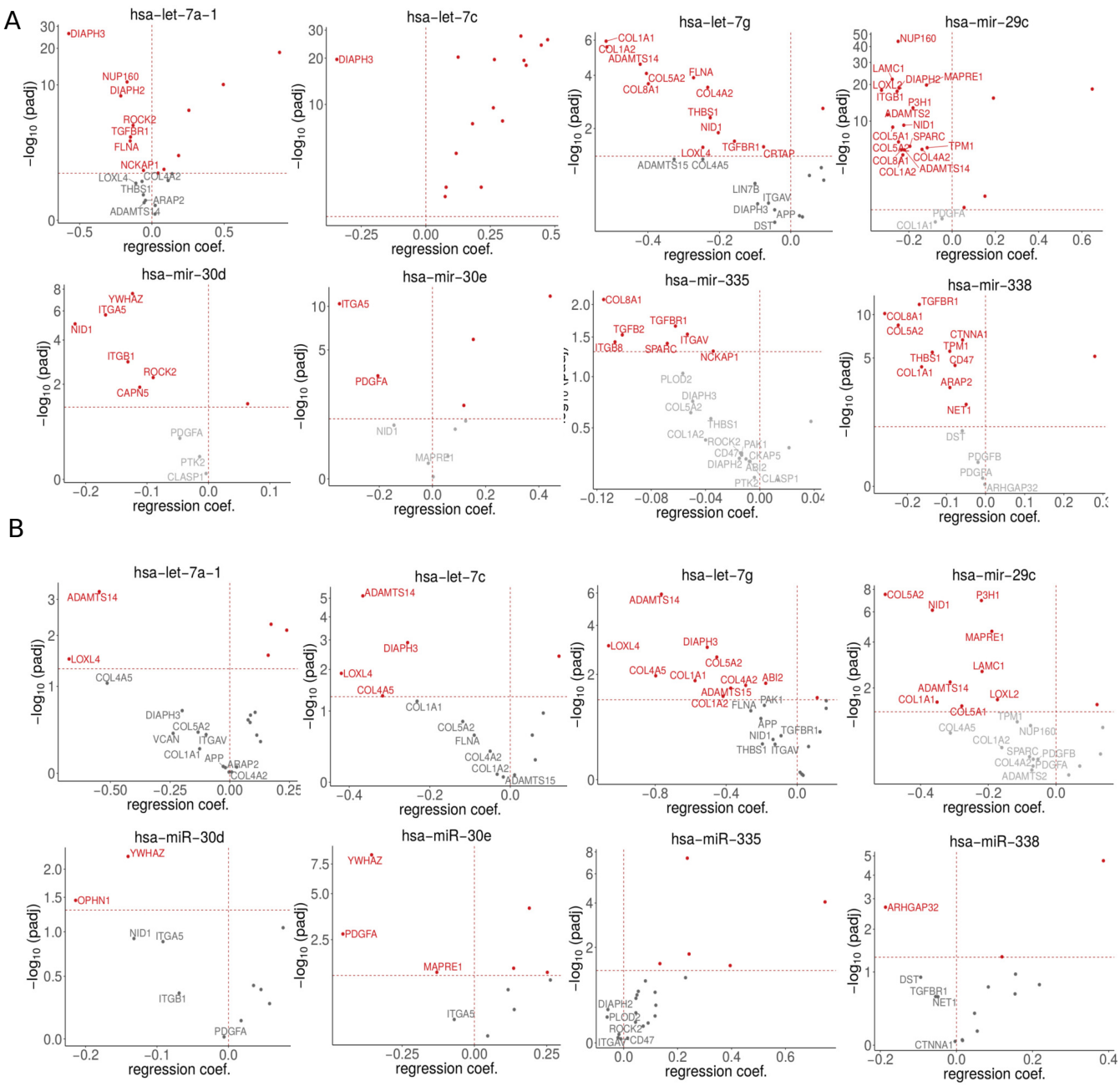

Fig. S8. miRNA:mRNA pairs in human BRCA and HCC show different degrees of association in comparison to other fibrosis-facilitated carcinomas. (A) Each plot shows recurrence of expression association of the particular miRNA l and its ECM-related target mRNA in human invasive breast carcinomas defined by regression coefficient β_l and corresponding p_{adj} value. The relationship of miRNA to individual mRNA was evaluated using the multivariant linear model which factors in: miRNA and mRNA expression, changes in DNA copy number (CNV) and promoter methylation status of the protein-coding genes. CNV and methylation data are used to assess the influence of miRNA-unrelated gene regulation. (B) Each plot shows recurrence of expression association of the particular miRNA and its ECM-related target mRNA in human HCCs defined by regression coefficient and corresponding p_{adj} value. Target genes having a p_{adj} value of miRNA:mRNA association assessment of ≤ 0.05 are represented in red. Gene names are displayed if their regression coef. β_l has negative value. * p-value ≤ 0.05, **p-value ≤ 0.01, *** p-value ≤ 0.001.

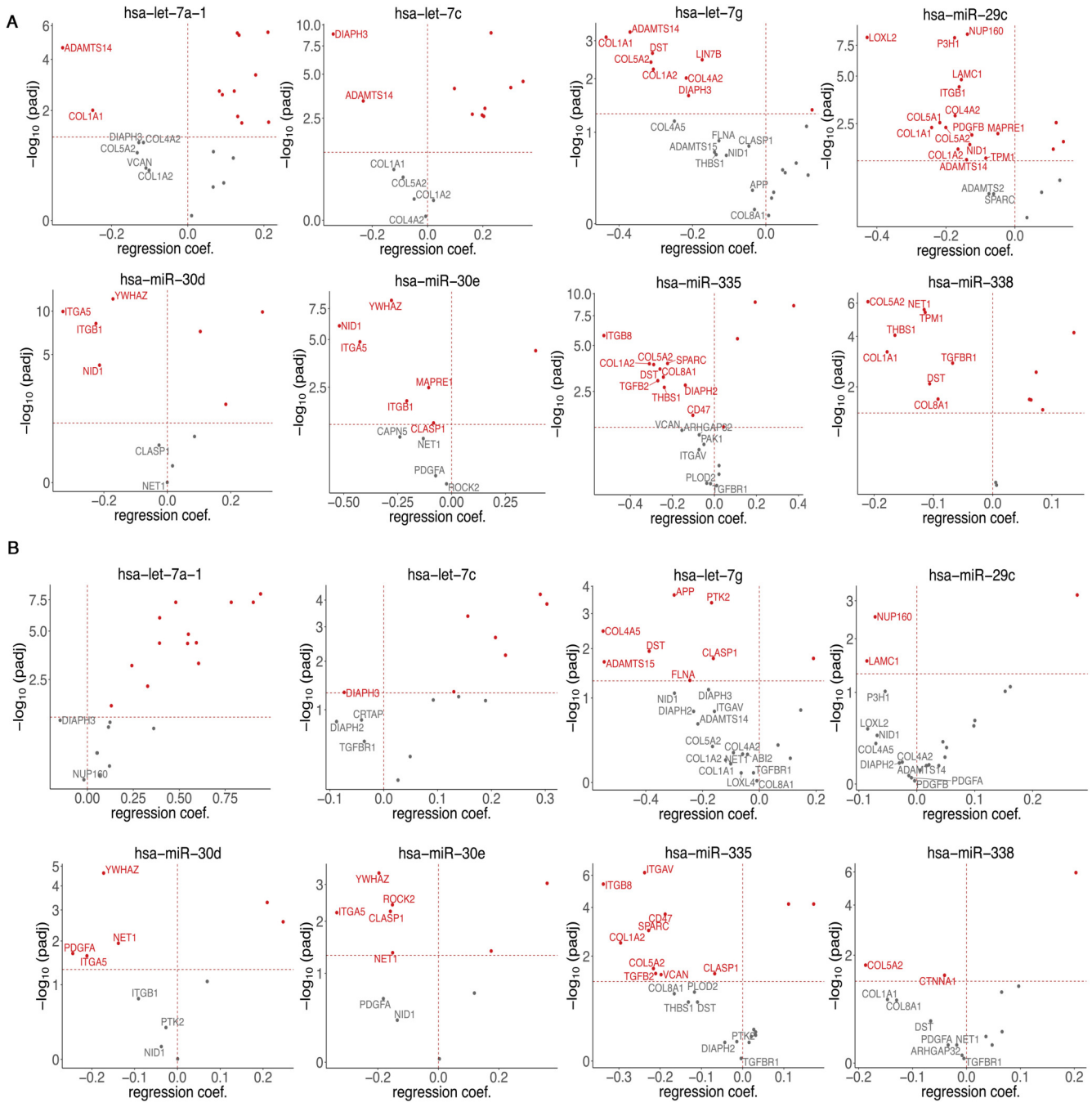

Fig. S9. miRNA:mRNA pairs in LUAD and LUSC show difference degrees of association in comparison to other fibrosis-facilitated carcinomas. (A-B) Plots show recurrence of expression association of the particular miRNA and its target mRNA in (A) human lung adenocarcinoma and (B) human lung squamous cell carcinoma defined by regression coefficient and corresponding p_{adj} value. Target genes having a p_{adj} value of miRNA:mRNA association assessment of ≤ 0.05 are represented in red. Gene names are displayed if their regression coef. has a negative value. * p-value ≤ 0.05, **p-value ≤ 0.01, *** p-value ≤ 0.001.

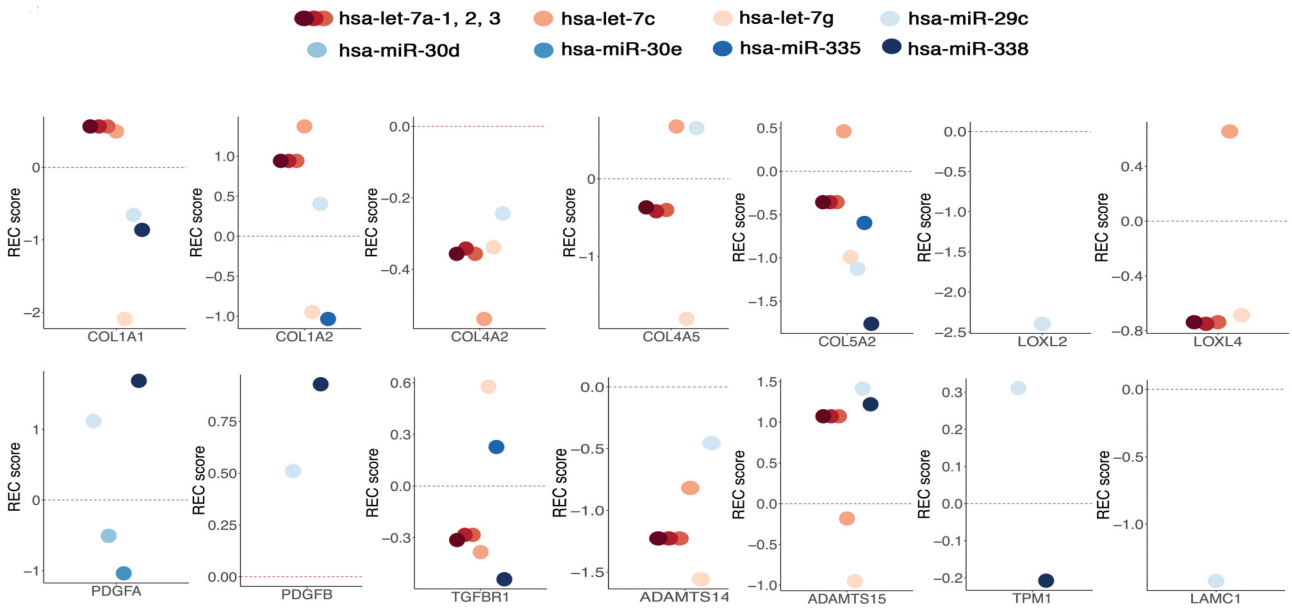

Fig. S10. The miRNA:mRNA pairs show different degrees of association in human fibrosis-facilitated carcinomas. Inferred association recurrence (REC) scores of miRNA:mRNA pairs depicted in Figure 1C in fibrosis-facilitated cancers (hepatocellular carcinoma (HCC), invasive breast carcinoma (BRCA), lung adenocarcinoma (LUAD) and lung squamous cell carcinoma(LUSC)). The sign of the REC score indicates the nature of the association, while its magnitude captures the recurrence consistency.

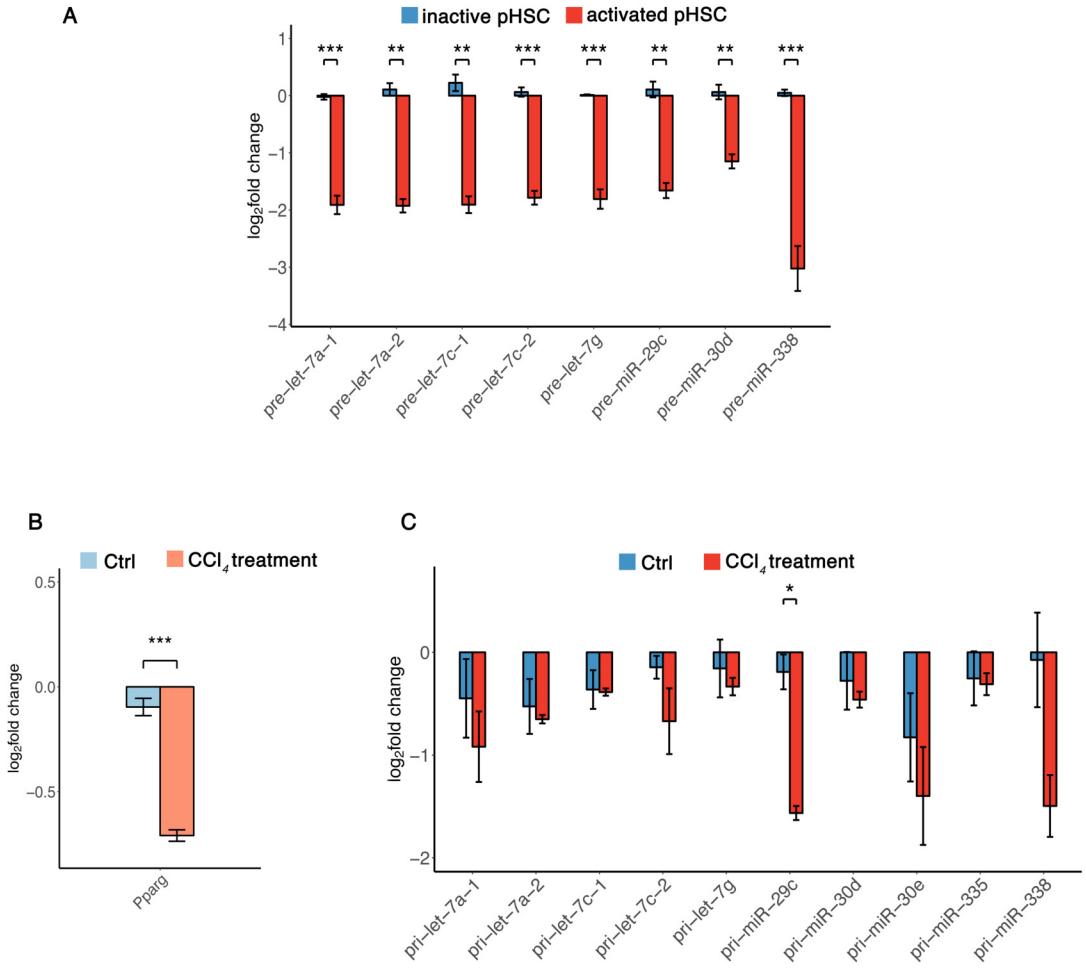

Fig. S11. Anti-fibrotic pre-miRNAs are downregulated in activated pHSC, as well as pri-miRNA in the murine CCl_4 model. (A) Relative expression of pre-miRNA in inactive and activated pHSCs. (B-C) Relative expression of: (B) *Pparg* and (C) pri-miRNAs in samples of CCl_4-treated mice in comparison to controls (mineral oil treatment). Data are shown as mean and standard error of the mean. * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

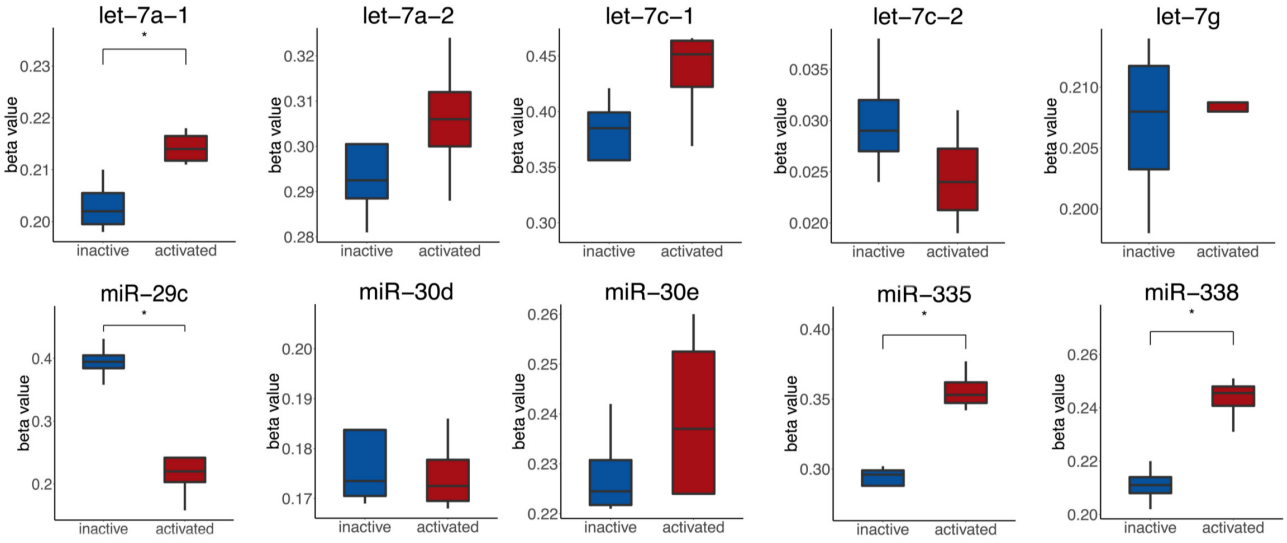

Fig. S12. DNA methylation of miRNA-encoding gene promoters contributes to regulation of expression of AF-miRNAs. Differential CpG DNA methylation of promoters of AF-miRNA-encoding genes in inactive and activated pHSCs. Data are shown as median, first and third quartile ("box") and 95% confidence interval of median ("whiskers"). * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

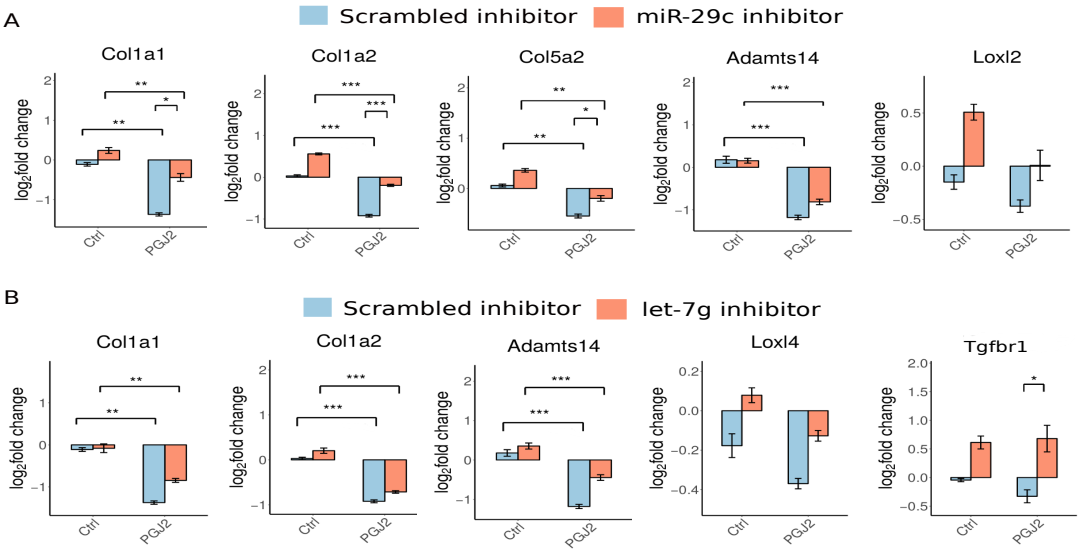

Fig. S13. Pparγ-mediated expression of miR-29c and let-7g inhibits fibrosis-associated genes. Relative expression of fibrosis-associated target genes of (A) miR-29c and (B) let-7g upon treatment of stable Pparγ-overexpressing GRX cells with the Pparγ agonist PGJ_2 in the presence of scrambled inhibitor (blue) or miRNA inhibitor (red). Data are shown as mean and standard error of the mean. * p-value $\leq$ 0.05, ** p-value $\leq$ 0.01, *** p-value $\leq$ 0.001.

Table S1. List of samples used in omics analyses and their corresponding metadata

Litter	Sample	Age/weeks	Gender	Liver-to-body-ratio	Sample type	Experiment
4	4-2C	32	F	4.8	control	sRNA-seq, RNA-seq
4	4-1N	32	F	26.3	nodule	sRNA-seq, RNA-seq
4	4-1T2	32	F	26.3	HCC	sRNA-seq, RNA-seq
25	25-5C	32	F	5.1	control	sRNA-seq, RNA-seq
25	25-4N	32	F	30.1	nodule	sRNA-seq, RNA-seq
25	25-4T1	32	F	30.1	HCC	sRNA-seq, RNA-seq
26	26-4C	31	F	4.2	control	sRNA-seq, RNA-seq
26	26-3N	31	F	26.8	nodule	sRNA-seq
26	26-3T2	31	F	26.8	HCC	sRNA-seq, RNA-seq
49	49-1C	52	M	4.9	control	RNA-seq
49	49-2T3	52	M	16.4	HCC	RNA-seq
2	2-1C	32	M	5	control	sRNA-seq
2	2-2N	32	M	24.3	nodule	sRNA-seq, RNA-seq
2	2-2T	32	M	24.3	HCC	sRNA-seq, RNA-seq
59	59-1N	30	M	23.4	nodule	RNA-seq
59	59-1T1	30	M	23.4	HCC	RNA-seq
59	59-1T2	30	M	23.4	HCC	RNA-seq
59	59-1T3	30	M	23.4	HCC	RNA-seq

Table S2. List of miRNA:mRNA pairs for which targeting was experimentally validated

miRNA	Target mRNA	Experiment
mmu-miR-29c-3p	Col1a1	Luciferase assay, qPCR
mmu-miR-29c-3p	Pdgfa	Luciferase assay, qPCR
mmu-miR-29c-3p	Pdgfb	qPCR
mmu-miR-29c-3p	Tpm1	qPCR
mmu-miR-29c-3p	Col1a2	qPCR
mmu-miR-29c-3p	Col4a2	qPCR
mmu-miR-29c-3p	Col4a5	qPCR
mmu-miR-29c-3p	Col5a2	qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Col1a1	Luciferase assay, qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Adamts15	Luciferase assay, qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Tgfbr1	Luciferase assay, qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Col4a2	qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Col4a5	qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Col5a2	qPCR
mmu-let-7a-5p, mmu-let-7c-5p, mmu-let-7g-5p	Col1a2	qPCR
mmu-miR-338-3p	Col1a1	Luciferase assay
mmu-miR-338-3p	Adamts15	Luciferase assay

Table S3. List of the miRNA:mRNA interactions, assays in which the interactions where validated and REC scores

Gene	miRNA	luciferase	qPCR	REC score
Pdgfa	miR-29c	y (validated)	y (validated)	positive
Adamts14	miR-29c	n	y (validated)	negative
Col1a1	miR-29c	y (validated)	y (validated)	negative
Col1a2	miR-29c	n	y (validated)	positive
Col4a2	miR-29c	n	y (validated)	negative
Col5a2	miR-29c	n	y (validated)	negative
Lamc1	miR-29c	n	y (validated)	negative
*Tpm*1	miR-29c	n	y (validated)	positive
Adamts15	miR-29c	y (validated)	n	positive
Tgfbr1	let-7g	y (validated)	y (validated)	positive
Col1a1	let-7g	y (validated)	y (validated)	negative
Adamts15	let-7g	y (validated)	n	negative
Adamts14	let-7g	n	y (validated)	negative
Col1a2	let-7g	n	y (validated)	negative
Coll4a2	let-7g	n	y (validated)	negative
Col4a5	let-7g	n	y (validated)	negative
Col5a2	let-7g	n	y	negative
Loxl4	let-7g	n	y (validated)	negative
Adamts15	let-7a	y (validated)	n	positive
Col1a1	let-7a	y (validated)	n	positive
Tgfbr1	let-7a	y (validated)	n	negative
Col1a1	let-7c	y (validated)	y (validated)	positive
Adamts15	let-7c	y (validated)	n	negative
Tgfbr1	let-7c	y (validated)	y (validated)	negative
Col1a2	let-7c	n	y (validated)	positive
Col4a2	let-7c	n	y (validated)	negative
Col4a5	let-7c	n	y (validated)	positive
Col5a2	let-7c	n	y (validated)	positive
Loxl4	let-7c	n	y (validated)	positive
Adamts15	miR-338	y (validated)	n	positive

Table S4. List of used vectors and their description

Vector	Vector description
pMSCV-Lin28A	Expression of murine Lin28A
pmirGLO	pmirGLO Dual-Luciferase miRNA Target Expression Vector
pMSCV	empty pMSCV-Lin28A vector (without insert)
pmirGLO-Col1a1	Expression of Col1a1-3'UTR with predicted let-7 and miR-29c target site in pmirGLO
pmirGLO-Pdgfa	Expression of Pdgfa1-3'UTR with predicted miR-29c target site in pmirGLO
pmirGLO-Adamts15	Expression of Adamts15-3'UTR with predicted let-7 and miR-338 target site in pmirGLO
pmirGLO-Tgfbr1	Expression of Tgfbr1-3'UTR with predicted let-7 target site in pmirGLO
pmirGLO-Col1a1-mut	Expression of Col1a1-3'UTR with mutated predicted let-7 and miR-29c target site in pmirGLO
pmirGLO-Pdgfa-mut	Expression of Pdgfa1-3'UTR with mutated predicted miR-29c target site in pmirGLO
pmirGLO-Adamts15-mut	Expression of Adamts15-3'UTR with mutated predicted let-7 and miR-338 target site in pmirGLO
pmirGLO-Tgfbr1-mut	Expression of Tgfbr1-3'UTR with mutated predicted let-7 target site in pmirGLO
pSV Sport PPAR gamma 1	Expression of mouse Pparg1
pcDNA3-Egr1	Expression of mouse Egr1

Additional data table S1 (dataset_one.xlsx)
Summary of statistical analysis. Related to the figures as listed in the table.

Additional data table S2 (dataset_two.xlsx)
List of differential expressed sRNAs and genes in SRF-VP16-driven nodular, tumor and corresponding control samples. (Sheet 1) List of differential expressed sRNAs obtained by sRNA-seq data analysis. List contains Gene ID information of all detected sRNAs, $\log_2$ fold changes in nodules and tumors compared to controls and corresponding p- and p_{adj}-values. (Sheet 2) Normalized read counts ($\log_2$ transformed) of sRNA in all samples. Read were produced using varianceStabilizingTransformation function of DESeq2 package. (Sheet 3) List of differential expressed genes obtained by RNA-seq data analysis. List contains Gene ID information of all detected mRNAs, $\log_2$ fold changes in nodules and tumors compared to controls and corresponding p- and p_{adj}-values. (Sheet 4) Normalized read counts ($\log_2$ transformed) of mRNAs in all samples. Reads were produced using varianceStabilizingTransformation function of DESeq2 package.

Additional data table S3 (dataset_three.xlsx)
REC scores of linear regression analyses. (Sheet 1) REC scores of miRNA:mRNA pairs and the corresponding p_{adj}-values. (Sheet 2-5) Regression coef. of miRNA:mRNA pairs and the corresponding p_{adj}-values for miRNAs expression, CNV and methylation values in BRCA (Sheet 2), LIHC (Sheet 3), LUAD (Sheet 4) and LUSC (Sheet 5).

Additional data table S4 (dataset_four.xlsx)
List of used primers, their sequence and targeted gene/miRNA. (Sheet 1) List of used primers, their sequence and targeted gene/miRNA. (Sheet 2) Sizes and genomic coordinates of amplicons analyzed in miRNA methylation study.

Additional data table S5 (dataset_five.xslx)
AF-miRNA and fibrosis-associated gene (mRNA) pairs conserved across patients of the TCGA cohort.

References

1. S Dalton, R Treisman, Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell* **68**, 597–612 (1992).
2. S Ohrnberger, et al., Dysregulated serum response factor triggers formation of hepatocellular carcinoma. *Hepatology* **61**, 979–989 (2015).
3. D Scholten, J Trebicka, C Liedtke, R Weiskirchen, The carbon tetrachloride model in mice. *Lab. Anim.* **49**, 4–11 (2015).
4. S Weiskirchen, CG Tag, S Sauer-Lehnen, F Tacke, R Weiskirchen, Isolation and Culture of Primary Murine Hepatic Stellate Cells. *Methods Mol. Biol.* **1627**, 165–191 (2017).
5. R Borojevic, et al., Establishment of a continuous cell line from fibrotic schistosomal granulomas in mice livers. *In Vitro Cell. Dev. Biol.* **21**, 382–390 (1985).
6. JL Jainchill, SA Aaronson, GJ Todaro, Murine sarcoma and leukemia viruses: assay using clonal lines of contact-inhibited mouse cells. *J. Virol.* **4**, 549–553 (1969).
7. U Negmadjanov, et al., TGF-β-mediated differentiation of fibroblasts is associated with increased mitochondrial content and cellular respiration. *PLoS ONE* **10**, e0123046 (2015).
8. A Dobin, et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
9. MI Love, W Huber, S Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
10. JM Ruijter, et al., Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. *Methods* **59**, 32–46 (2013).
11. B Daniel, BL Balint, ZS Nagy, L Nagy, Mapping the genomic binding sites of the activated retinoid X receptor in murine bone marrow-derived macrophages using chromatin immunoprecipitation sequencing. *Methods Mol. Biol.* **1204**, 15–24 (2014).
12. M Ehrich, et al., Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15785–15790 (2005).
13. SR Viswanathan, et al., Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41**, 843–848 (2009).
14. P Tontonoz, E Hu, RA Graves, AI Budavari, BM Spiegelman, mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. *Genes Dev.* **8**, 1224–1234 (1994).
15. M Reczko, M Maragkakis, P Alexiou, I Grosse, AG Hatzigeorgiou, Functional microRNA targets in protein coding sequences. *Bioinformatics* **28**, 771–776 (2012).
16. RC Friedman, KK Farh, CB Burge, DP Bartel, Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
17. LJ Core, JJ Waterfall, JT Lis, Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).

18. AA Sigova, et al., Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2876–2881 (2013).
19. X Luo, M Chae, R Krishnakumar, CG Danko, WL Kraus, Dynamic reorganization of the AC16 cardiomyocyte transcriptome in response to TNFÎ± signaling revealed by integrated genomic analyses. *BMC Genomics* **15**, 155 (2014).
20. A Wang, et al., Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**, 386–399 (2015).
21. C Estaras, C Benner, KA Jones, SMADs and YAP compete to control elongation of Î²-catenin:LEF-1-recruited RNAPII during hESC differentiation. *Mol. Cell* **58**, 780–793 (2015).
22. IM Min, et al., Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* **25**, 742–754 (2011).
23. B Fang, et al., Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo. *Cell* **159**, 1140–1152 (2014).
24. RS Illingworth, JJ Holzenspies, FV Roske, WA Bickmore, JM Brickman, Polycomb enables primitive endoderm lineage priming in embryonic stem cells. *Elife* **5** (2016).
25. AA Sigova, et al., Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**, 978–981 (2015).
26. GA Busslinger, et al., Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503–507 (2017).
27. Y Zhang, et al., The hepatic circadian clock fine-tunes the lipogenic response to feeding through RORÎ±/Î³. *Genes Dev.* **31**, 1202–1211 (2017).
28. S Heinz, et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
29. CE Grant, TL Bailey, WS Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
30. A Sandelin, W Alkema, P Engstrom, WW Wasserman, B Lenhard, JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91–94 (2004).
31. JY Chen, et al., Tricyclic Antidepressants Promote Ceramide Accumulation to Regulate Collagen Production in Human Hepatic Stellate Cells. *Sci Rep* **7**, 44867 (2017).
32. D Ohlund, et al., Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).
33. A Jacobsen, et al., Analysis of microRNA-target interactions across diverse cancer types. *Nat. Struct. Mol. Biol.* **20**, 1325–1332 (2013).