

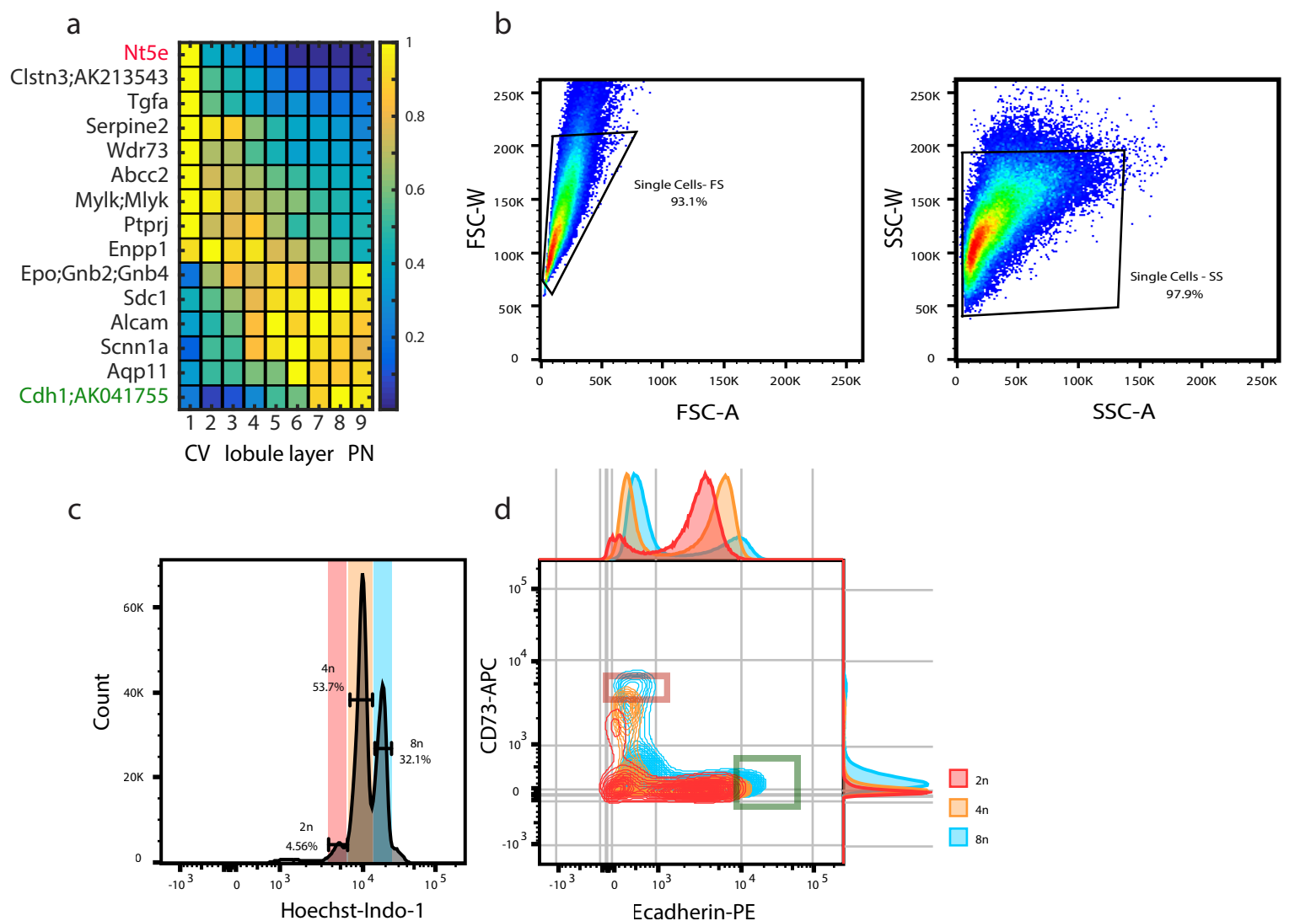
In the format provided by the authors and unedited.

Spatial sorting enables comprehensive characterization of liver zonation

Shani Ben-Moshe^{1,3}, Yonatan Shapira^{1,3}, Andreas E. Moor ^{1,2}, Rita Manco¹, Tamar Veg¹,
Keren Bahar Halpern¹ and Shalev Itzkovitz ^{1*}

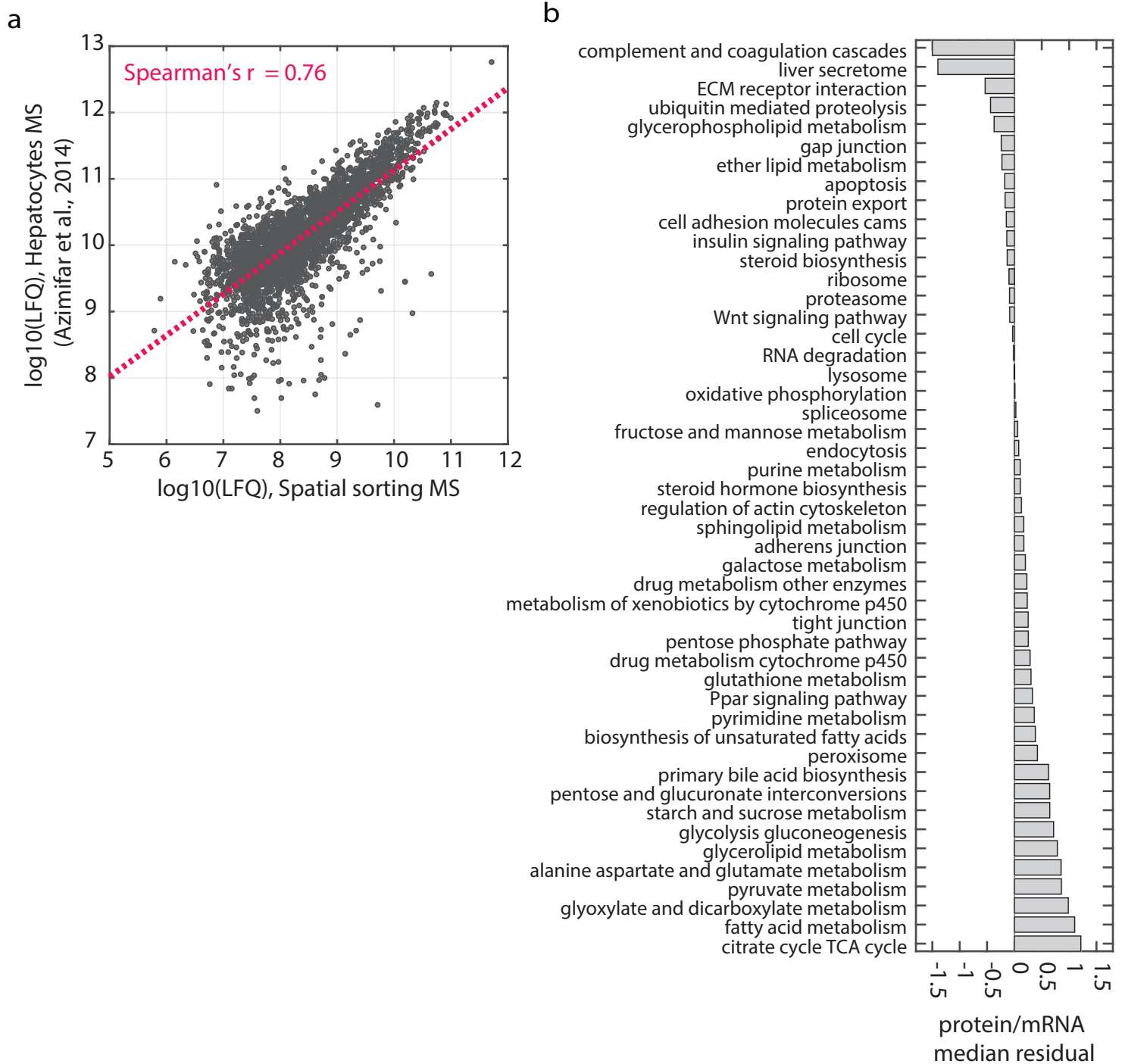
¹Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. ²Present address: Institute of Molecular Cancer Research, University of Zurich, Zurich, Switzerland. ³These authors contributed equally: Shani Ben-Moshe, Yonatan Shapira. *e-mail: shalev.itzkovitz@weizmann.ac.il

Supplementary Information

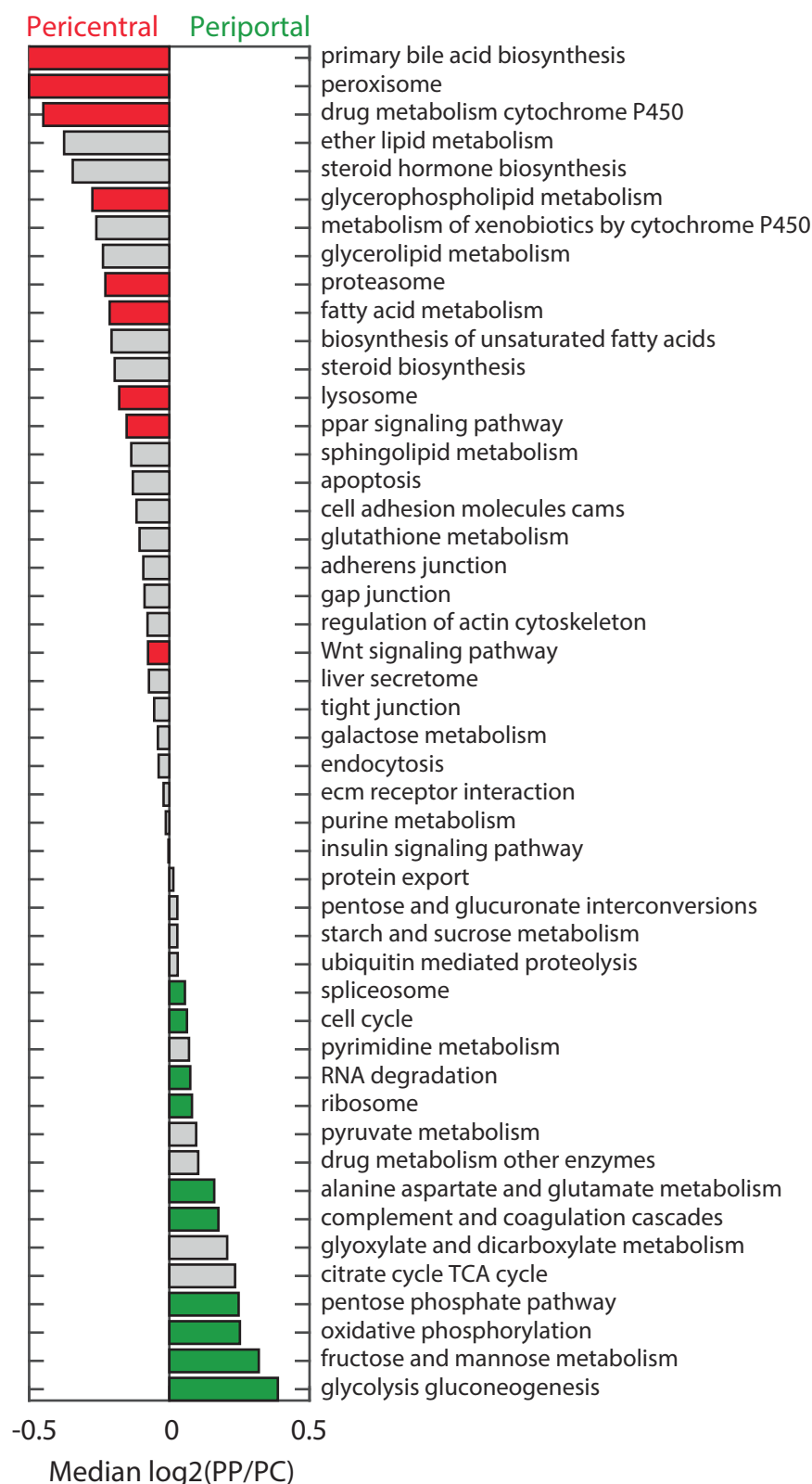


Supplementary Fig. 1 | Surface markers candidates for spatial sorting and ploidy gating in FACS.

a, Surface markers zonation profiles. Data from 7. Expression is normalized to the maximal level for each gene. Genes are sorted by expression center of mass. b, FACS gates used to enrich for single hepatocytes. c-d, Gating by ploidy levels. c, Histogram of cells according to Hoechst stains, proportional to DNA content, each peak represents a different ploidy class with different proportions in the hepatocytes population. Red – diploid cells, orange – tetraploid cells, blue – octoploid cells. d, distribution of hepatocytes according to E-cadherin and CD73 intensities, stratified by ploidy class. Red and green rectangles are 1 and 8 gates selected for the 4n hepatocytes.

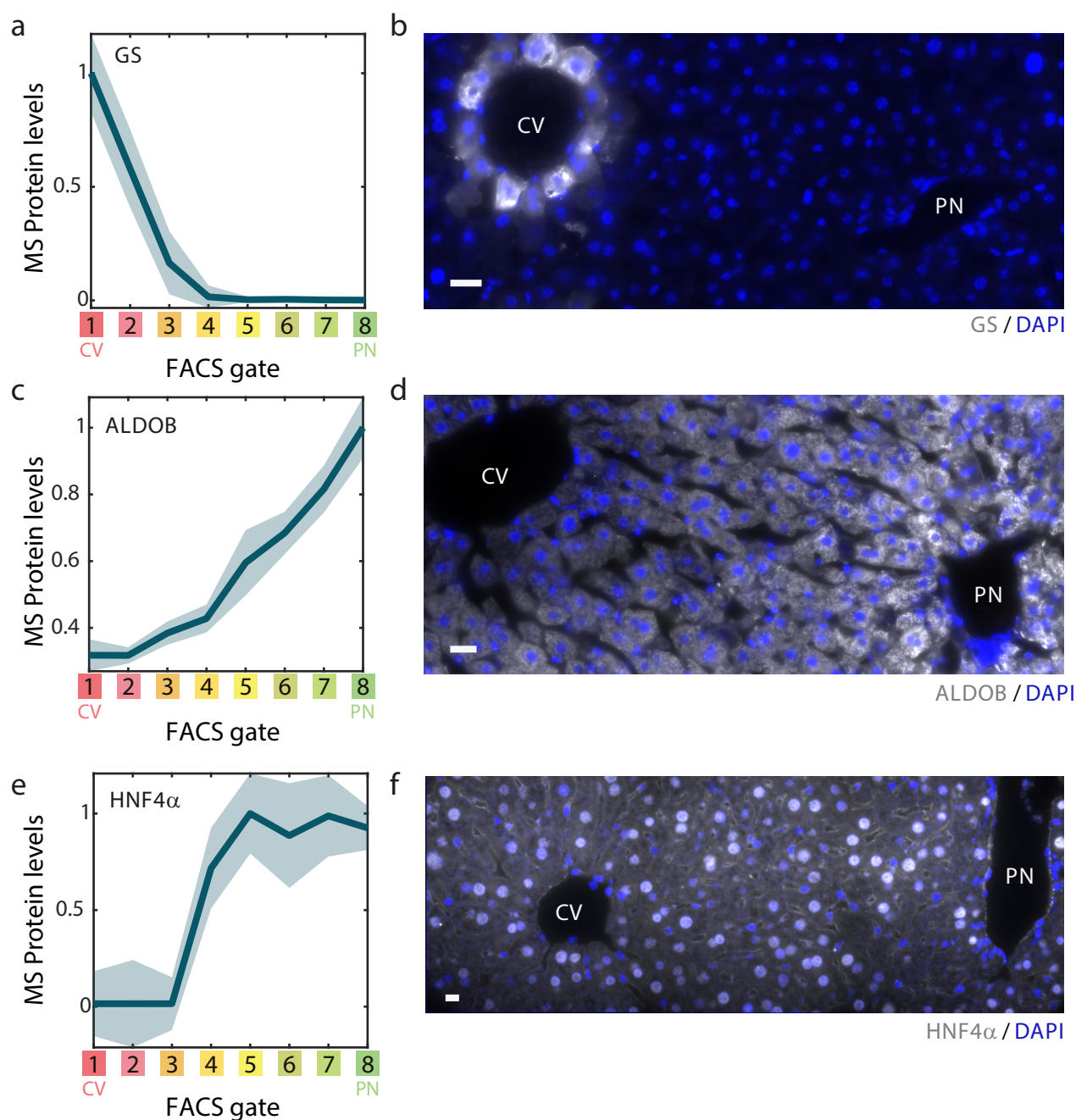


Supplementary Fig. 2 | Comparisons of proteomic dataset. a, Scatter plot of log10(LFQ) of mean proteins over all FACS gates for 5 mice and log10(LFQ) of previously published mouse hepatocytes mass-spec measurements¹⁴. Red dashed line marks the regression line, Spearman's $r=0.76$. Number of proteins = 2831. b, Protein to mRNA residuals from regression line, bar represent median of genes belonging to the KEGG pathway. See main Fig. 4.

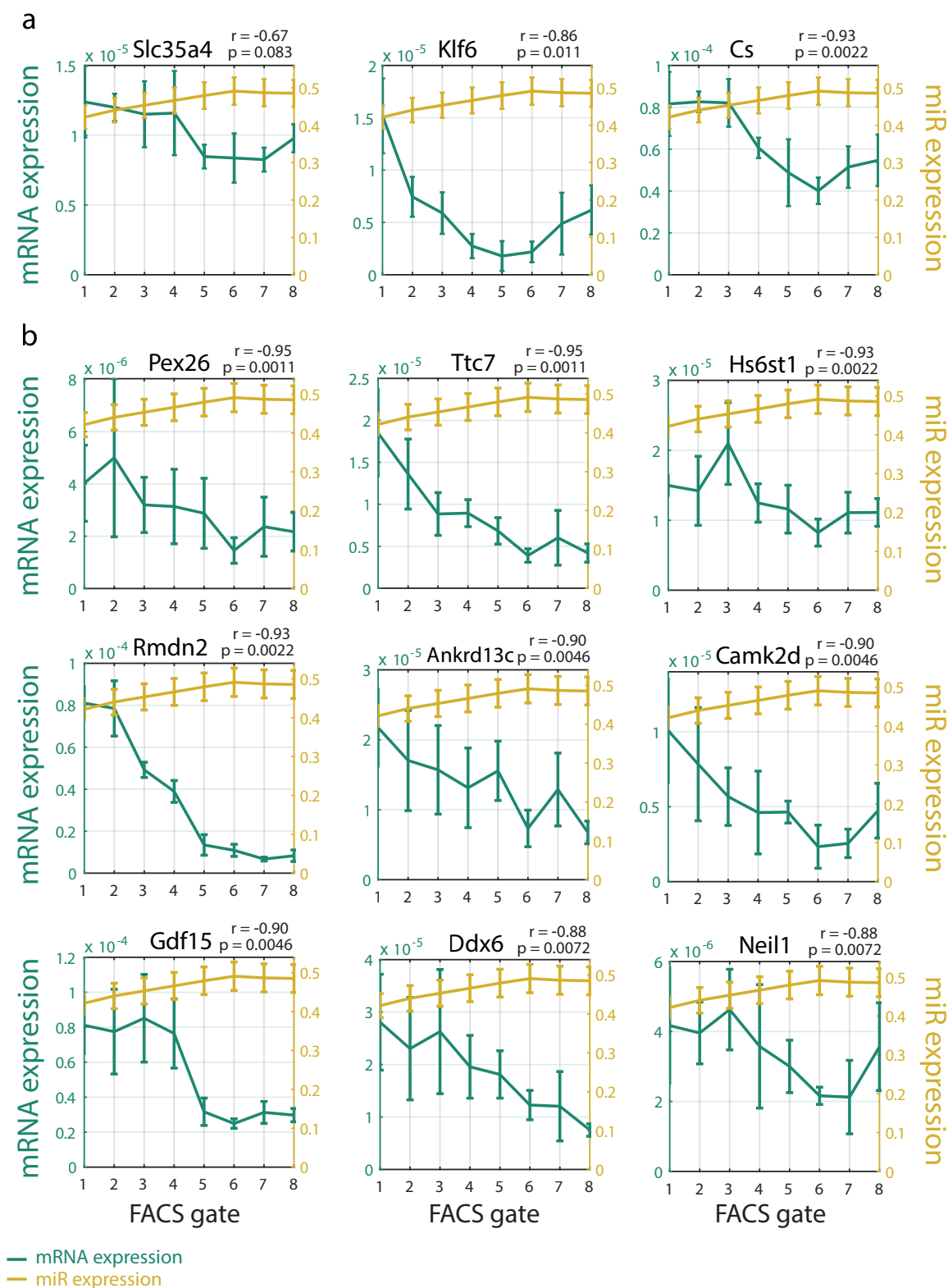


Supplementary Fig. 3 | Zonation of hepatocyte proteome by KEGG pathways.

Median log2 of ratio between periportal and pericentral gates of proteins belonging to the selected KEGG families over the 5 mice. Colored bars represent KEGG pathways with FDR < 0.2 of p-values obtained from two-sided single sample sign rank test for each of the KEGG pathways. Red represent negative log2 ratio, indicating pericentral enrichment, and green represent positive values, indicating periportal enrichment.



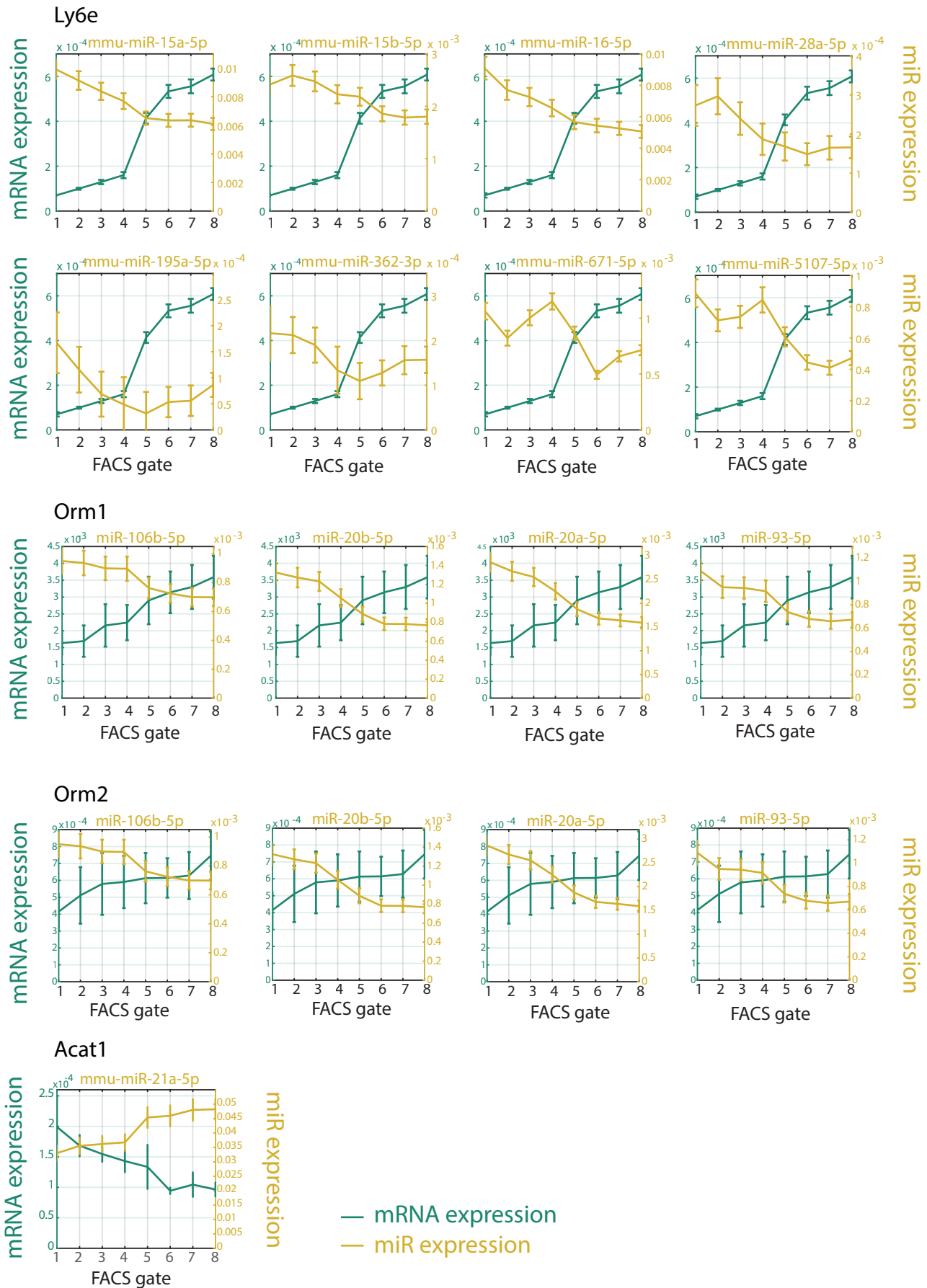
Supplementary Fig. 4 | Immuno-Fluorescence of selected proteins validates proteins zonation. Protein levels of three proteins - Glutamine synthetase (GS) (a-b), Aldolase b (ALDOB) (c-d) and HNF4 α (e-f). Graphs (a,c,e) presenting protein levels according to mass-spectrometry (MS) results in the eight zoned FACS gates populations. Protein levels are normalized to the maximal expression level of each protein across the eight populations. N=5 mice. Lines indicate means and patches indicate SEM. Images (b,d,f) are representative IF lobule scans showing gradients of proteins (white) along the protocentral axis. Blue – DAPI. CV – central vein, PN – portal node, scale bar – 10 μ m. IF experiments were performed on two mice in total for GS and ALDOB with similar results and three mice for HNF4 α with similar results.



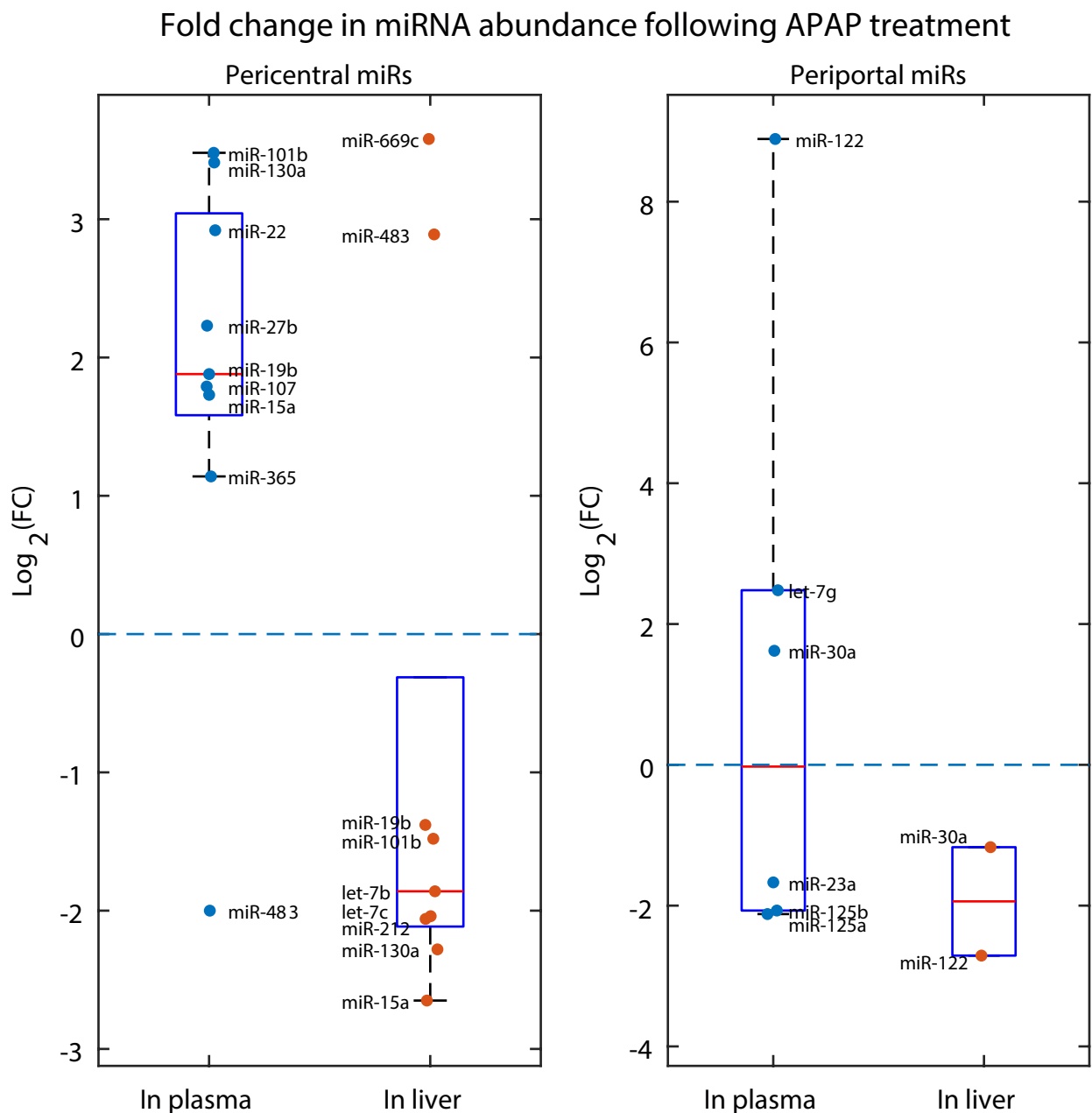
Supplementary Fig. 5 | Selected anti-correlated target genes of miR-122-5p.

a, miR-122-5p was found to be anti-correlated with some of its canonical targets, *Slc35a4*, *Klf6* and *Cs* (following 29). b, Top nine anti-correlated predicted targets of miR-122-5p.

Expression is normalized to the sum. Lines represent mean of 5 mice for mRNA (green) and 3 mice for miRNA (yellow). Error bars are SEM. Shown are Spearman correlations and their respective p-values, computed using exact permutation distributions.

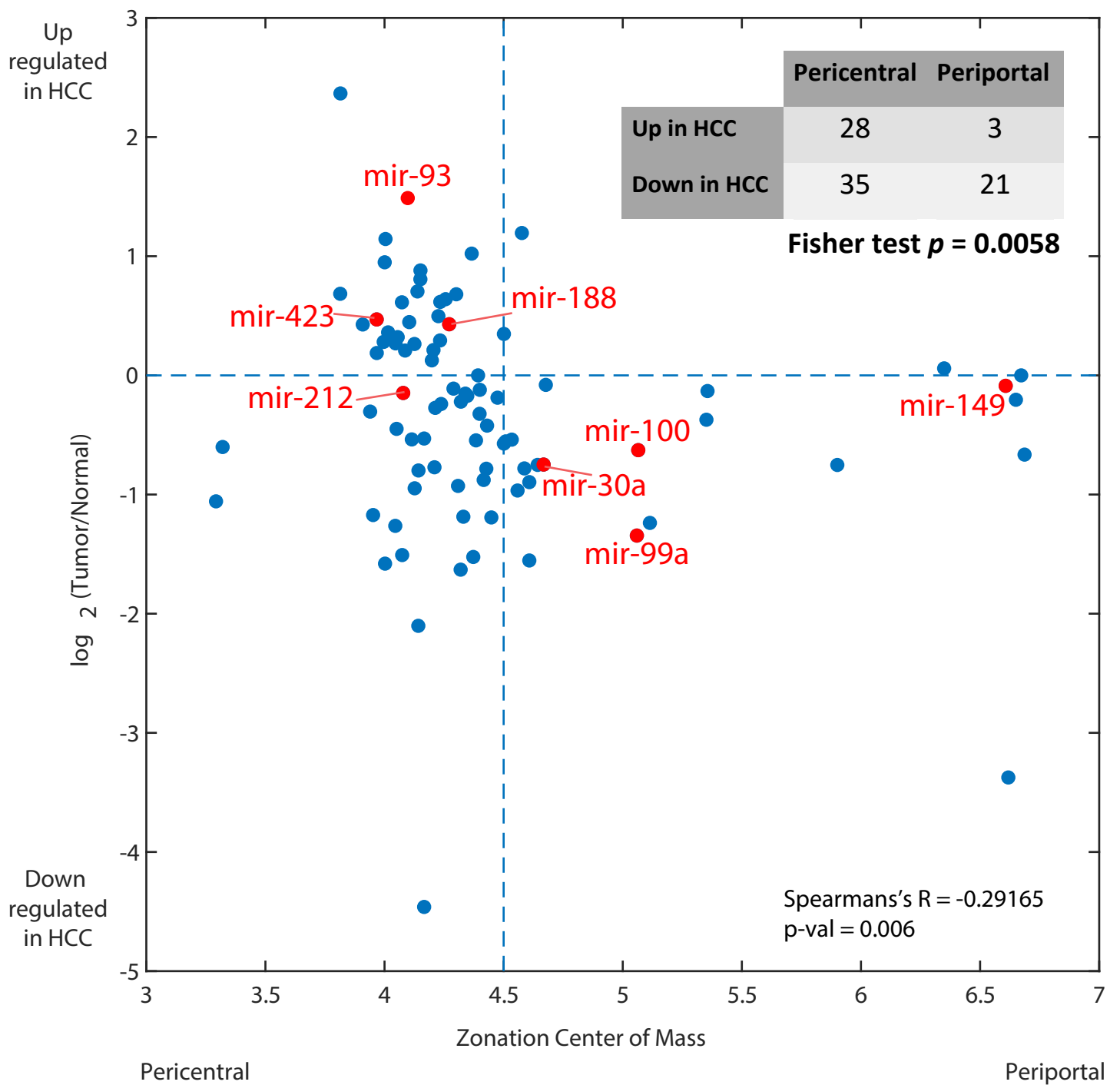


Supplementary Fig. 6| Target genes significantly anti-correlated with their potential regulating miRNAs. Expression is normalized to the sum. Lines represent mean of 5 mice for mRNA (green) and 3 mice for miRNA (yellow). Error bars are SEM.



Supplementary Fig. 7 | Pericentrally-zonated miRNA are enriched in the plasma of mice undergoing pericentral damage via acetaminophen (APAP) overdose.

Shown are fold-changes of miRNA levels in blood plasma and liver of APAP-treated in mice, relative to controls, stratified by zonation. Pericentral miRNA have zonation center of mass smaller than 4.5 and zonation FDR ≤ 0.2 , periportal miRNA have zonation center of mass larger than 4.5 and zonation FDR ≤ 0.2 . miRNA levels in APAP-treated mice were obtained from Wang et al. (2009)⁴⁵, n=4 mice. a, Pericentral miRNAs are more abundant in plasma and less abundant in liver of APAP-treated mice (two-tailed Wilcoxon $p = 0.050$). b, Periportal miRNAs are not significantly more abundant in plasma compared to liver tissue (two-tailed Wilcoxon $p = 0.43$). Horizontal red lines are medians, boxes denote 25-75 percentiles, whiskers are 1.5x inter quartile range.



Supplementary Fig. 8 | Periportally-zonated miRNA are down-regulated in hepatocellular carcinoma. The log₂ fold change of miRNAs in HCC tumors (n = 372 TCGA tumor samples) compared to normal samples (n = 50 TCGA normal samples) is plotted against the miRNAs' expression center of mass. Dashed line mark no zonation in the x-axis and no change in tumors in the y-axis. Red – miRNAs with predicted Wnt component targets, found in TargetScan. P-value of Spearman correlation was computed using exact permutation distributions, with a two-tailed hypothesis.

Supplementary Tables

Supplementary Table 1 – Expression of Non-hepatocytes genes in RNA extracted from the FACS gates

Table includes 900 genes specifically expressed in one of the cell types in the liver. First nine values are the relative expression of the genes in the different cell types. Data for endothel, Kupffer, Liver Capsule Macrophages (LCMs), T cells, B cells, Neutrophils, plasmacytoid dendritic cells (pDCs) and hepatocytes are taken from Haplern et al.¹³. Cholangiocytes expression data is taken from Li et al.⁷⁰. Next, the relative expression of these genes in each FACS gate population is presented, in RNA levels and protein levels. na – not detected in proteomics data (see Methods).

Supplementary Table 2 - mRNA UMI counts

Eight sorted populations (1-8) from five different mice (m1-m5) were collected and mRNA libraries were prepared according to mcSCRB-seq protocol⁶⁵, and zUMIs pipeline⁶⁶. Exonic UMI counts were normalized using TMM⁶⁷ implemented in EdgeR⁶⁸. The table values are the normalized CPM (UMI counts per million). Note that in our analyses, we have discarded m2_2 and m3_5 as their number of reads were less than 200K.

Supplementary Table 3 - Proteins MS/MS quantification

Eight sorted populations (1-8) from five different mice (m1-m5) were collected and extracted proteins were processed by LC-MS/MS. Data were analyzed and quantified with MaxQuant 1.5.2.8⁶⁹ against the mouse Uniprot database. Table Values are log₂(LFQ) – relative label-free quantification of proteins across samples. Table also shows sequence coverage, Razor+Unique peptides and MS/MS spectral count for each sample, as well as protein identifiers.

Supplementary Table 4 – mRNA and proteins of spatially-sorted hepatocytes

Shown are the data for 3,051 genes with matched transcriptomic and proteomic information. The mean and SEM of the five repeats in each FACS gate, normalized to their sample sum, were calculated for both mRNA and protein. Kruskal-Wallis p-value and FDR was assigned to each of the mRNAs/proteins.

Supplementary Table 5 – Zonation of differentially expressed genes in HNF4 α KO male mice compared to WT male mice

Table includes 1,166 hepatocyte genes found by Holloway et al.²¹ to be up/down regulated upon HNF4 α KO in male mice. Table shows the fold change in expression of KO/wt and the FDR q-val. Also shown is the zonation data from our spatial sorted hepatocyte populations – the gene center of mass, the zonation FDR q-val, and the relative RNA expression of the genes in each of the FACS gates.

Supplementary Table 6 – Zonation profiles of detected high-confidence miRs

Shown are the mean and SEM of the sum-normalized expression of miRNAs for each FACS gate, p-value for zonation (Kruskal-Wallis test), false discovery rate (FDR, using Benjamini-Hochberg procedure) and visualization of zonation profiles.

Supplementary Table 7 – Zonation profiles of high-confidence miRNA-target pairs

Includes expression level per FACS gate (as fraction of total) of miRNAs and their regulated genes, and Spearman correlation. Predicted interactions were obtained from TargetScan 7.2, with weighted context++ score percentile ≥ 95 .

Supplementary Table 8 – Zonation profiles of target genes and their cumulative miRNA profiles

Includes expression level per FACS gate of genes and their cumulative miRNA profiles, Spearman correlation, p-value for the interaction as obtained by the network analysis (see Methods), FDR and visualization of gene and cumulative miRNA zonation profiles. NA – interactions that have not been considered due to low expression of the target or regulating miRs (Methods).

Supplementary Table 9 – Mutation status in Wnt pathway genes and miRNA expression for all samples in the TCGA LIHC dataset

Supplementary Table 10 - Primer sequences used for miRNA qRT-PCR

Supplementary Table 11 - SRA sample table and zUMI barcodes

For each of the 40 mRNA samples sequenced, the table specifies which pool deposited in SRA includes the samples, the mcSCRB-seq barcode used for the sample, its respective accession number in SRA and the settings defined in zUMI pipeline for our analysis.