Spatial sorting enables comprehensive characterization of liver zonation

Shani Ben-Moshe1,3, Yonatan Shapira1,3, Andreas E. Moor1,2, Rita Manco1, Tamar Veg1, Keren Bahar Halpern1 and Shalev Itzkovitz1*

The mammalian liver is composed of repeating hexagonal units termed lobules. Spatially resolved single-cell transcriptomics has revealed that about half of hepatocyte genes are differentially expressed across the lobule, yet technical limitations have impeded reconstructing similar global spatial maps of other hepatocyte features. Here, we show how zonated surface markers can be used to sort hepatocytes from defined lobule zones with high spatial resolution. We apply transcriptomics, microRNA (miRNA) array measurements and mass spectrometry proteomics to reconstruct spatial atlases of multiple zonated features. We demonstrate that protein zonation largely overlaps with messenger RNA zonation, with the periportal Wntα receptor Wingless-related integration site (Wnt) receptors Fzd7 and Fzd8 and the pericentral Fzd8 receptors. This spatial heterogeneity of hepatocytes may also exist for other cellular features, including proteins, metabolites and regulatory molecules such as miRNAs. However, achieving similar global zonation maps for cellular features beyond mRNA has encountered technical difficulties.

Immunohistochemistry enables the measurement of protein levels with high spatial resolution but is low-throughput and often limited by lack of availability of antibodies. Laser capture microdissection and digitonin perfusion enable extracting large numbers of peripherally or pericentrally enriched cells3,4. However, these techniques are limited in spatial resolution. Single-cell measurements of cellular features beyond mRNA are starting to emerge5; however, these technologies are less mature in tissues. A methodology that establishes high spatial resolution would enable generating organ-wide spatial atlases of key features, such as methylation patterns, chromatin conformations, miRNA content and proteomics. In the liver, such measurements would broaden our understanding of the regulation of liver zonation and could be used to model liver metabolic function more precisely.

In this study, we developed an approach termed ‘spatial sorting’, which uses surface markers with discordant zonation profiles to isolate very large amounts of hepatocytes from defined lobule layers. We used these for high-throughput profiling of mRNAs, miRNAs and proteins (Fig. 1c), revealing previously unknown features of liver zonation. These include a comprehensive proteomic zonation map and the identification of zonated miRNA with discordantly zonated target genes. Our approach can be readily applied to profile other cellular features of hepatocytes and other cell types in health and disease.

Results

Spatial sorting enables isolating bulk hepatocyte populations from different lobule layers. We used our recently reconstructed mRNA zonation map6 to identify zonated surface markers with a large dynamic range in expression, spanning several radial lobule layers (Fig. 1a and Supplementary Fig. 1a). We argued that the combined staining of two inversely zonated surface proteins would be informative for inferring the lobule positions of single hepatocytes (Fig. 1b), which would facilitate cell sorting of many cells according to their spatial origin (Fig. 1b,c). CD73, encoded by the gene Nt5e, is an enzyme that converts mononucleotides to nucleosides and exhibits perilobular zonation. E-cadherin, a cell–cell adhesion glycoprotein encoded by Cdh1, exhibits perilobular zonation (Fig. 2a). We used immunofluorescence to validate the zonation of these two surface markers at the protein level (Fig. 2b,c).

We perfused the livers of five mice fed ad libitum to dissociate single cells and performed fluorescence-activated cell sorting (FACS) of isolated hepatocytes stained with antibodies against CD73 labelled with allophycocyanin (APC) and E-cadherin labelled with phycoerythrin (PE). We filtered hepatocytes by size and selected cells that were negative for the endothelial cell marker CD31 and...
the immune cell marker CD45, to avoid pairs of hepatocytes and non-parenchymal cells (NPCs). We further filtered out non-viable cells and selected tetraploid hepatocytes using Hoechst staining (Fig. 3a and Supplementary Fig. 1b). Stratifying hepatocytes by ploidy was important to obtain precise lobule localization (Supplementary Fig. 1c,d). The selected hepatocytes displayed strong anti-correlation in the fluorescence of CD73 and E-cadherin, as expected from the zonated expression patterns (Fig. 3b).

We defined eight gates based on the combined fluorescence of CD73 and E-cadherin (Fig. 3b). To ensure reproducibility, the gates were defined as the percentiles of the marginal expression levels of each surface marker, compared to the unstained control. To validate that our defined gates represented sequential lobule layers, we performed bulk RNA-sequencing (RNA-seq) on 10,000 sorted hepatocytes from each gate. Contamination of NPC RNA was negligible and uniform across all eight isolated populations, validating our isolation and sorting approach (Supplementary Table 1). We compared the zonation profiles obtained via spatial sorting to our spatially resolved single-cell RNA-seq map. Zonation profiles were highly concordant (Fig. 3c and Supplementary Table 2), demonstrating the feasibility of our approach for isolating bulk hepatocytes with high spatial resolution.

Mass spectrometry proteomic measurements of spatially sorted hepatocytes. We next applied spatial sorting to reconstruct the zonation patterns of the hepatocyte proteome. To this end, we sorted 100,000 hepatocytes from each of the eight FACS gates and performed mass spectrometry proteomics. For each mouse and

---

**Fig. 1** | Spatial sorting approach for isolating large amounts of hepatocytes from distinct layers with high resolution. a, Identification of zonated surface markers. b, FACS enables defining gates that enrich for zonated hepatocytes according to their surface marker expression. c, Spatially sorted hepatocytes can be measured using multiple assays that require large input material, such as the RNA-seq, mass spectrometry and miRNA microarray applied in the current study.

**Fig. 2** | CD73 and E-cadherin are inversely zonated surface markers. a, CD73, encoded by Nt5e, and E-cadherin, encoded by Cdh1, are surface markers that are zonated at the mRNA level (data taken from Halpern et al.). n = 1,415 cells from 3 mice. The lines show the sum-normalized mean of all cells; the shaded regions are ± s.e.m. b, CD73 and E-cadherin proteins are zonated. An example of a lobule stained by immunofluorescence with antibodies against CD73 (red) and E-cadherin (green) is shown. Blue, DAPI nuclear stain. Scale bar, 10 µm. The experiment was performed independently on three different mice. c, Quantification of immunofluorescence images (n = 8 lobules from three mice). The lines represent the mean intensity measured in the lobule layer; the shaded regions are ± s.e.m. across the eight lobules.
gate we also isolated 10,000 cells and applied bulk RNA-seq. The mass spectrometry measurements yielded 3,210 identified proteins (Supplementary Table 3). The hepatocyte protein content averaged over all FACS gates was highly correlated with previous bulk measurements\(^{14}\) (Spearman’s \(r = 0.75\); Supplementary Fig. 2a). Our dataset included 3,051 proteins with matched mRNA (Supplementary Table 4). The means over all gates of protein and mRNA levels were positively correlated (Spearman’s \(r = 0.5\), \(P = 1.2 \times 10^{-181}\), \(n = 3,051\)). Yet, for some proteins, there was a marked difference in protein and mRNA relative abundances (Fig. 4). These predominantly included proteins secreted by hepatocytes. For example, \(\text{Alb}\), which encodes the secreted carrier protein albumin, was the most highly abundant hepatocyte mRNA (0.050 ± 0.004 of cellular transcripts) but was ranked only 64 in protein levels (0.0034 ± 0.0002 of cellular proteins). Other secreted proteins, which were ranked substantially higher in mRNA compared to the protein level, included apolipoproteins encoded by \(\text{Apoa1}\), \(\text{Apoa2}\), \(\text{Apoe}\), \(\alpha\)-antitrypsin encoded by \(\text{Serpin}\) genes, complement system proteins and vitronectin, encoded by \(\text{Vtn}\) (Fig. 4a). A similar discordance between the levels of mRNA and proteins for secreted genes was previously observed in mammalian cell lines\(^{15}\). Ribosomal mRNA and proteins had a protein-to-mRNA ratio close to 1, whereas genes of the tricarboxylic acid cycle had substantially higher protein-to-mRNA levels (Fig. 4b and Supplementary Fig. 2b). \(\text{Cps1}\), which encodes the urea cycle

---

**Fig. 3** | Spatial sorting reliably captures the different lobule layers. a, FACS gating strategy. FSC-A and SSC-A were used for size selection of hepatocytes. Non-viable cells were filtered out using the Zombie Green Viability Kit. Staining with antibodies to CD31 and CD45 enabled the gating out of NPCs. Tetraploid hepatocytes were selected based on Hoechst staining. b, Distribution of the included cells (40–60% from all events) according to the intensities of CD73 and E-cadherin. The grey lines mark the unstained control limits; the rectangles and numbers mark the gates used for spatially sorted populations. The distributions from five independent mice were similar. c, Maximal level-normalized expression patterns of selected genes along the different FACS gates (blue, \(n = 5\) mice), compared with the interpolated maximal level-normalized zonation profiles based on Halpern et al.\(^{7}\) (yellow). The lines are the mean of each FACS gate for 5 mice (blue) and the mean of each interpolated layer for the 1,415 cell from 3 mice (yellow). The line patches represent the s.e.m.
enzyme carbamoyl phosphate synthase was ranked first in protein content (0.0682 ± 0.0066), but only 478 in mRNA expression (2.88 × 10^{-4} ± 5.4 × 10^{-1} of cellular transcripts; Fig. 4a). Thus, the relative expression levels of mRNA and proteins differ for distinct functional classes.

**Zonation patterns of the hepatocyte proteome.** We next examined whether the hepatocyte proteome exhibited zonated patterns. We found that 55% of hepatocyte proteins (1,672 out of 3,051) were significantly zonated (false discovery rate (FDR) < 0.05, Kruskal–Wallis test; Fig. 5a). Periportal and pericentral enrichment of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways largely recapitulated previous zonation studies performed by using RNA⁴. Bile acid biosynthesis, lipid metabolism and cytochrome P450 xenobiotic metabolism were pericentral zonated, while gluconeogenesis, oxidative phosphorylation and complement and coagulation cascades were periportal zonated (Supplementary Fig. 3). We validated the measured protein zonation using immunofluorescence for representative pericentral and periportal proteins (Supplementary Fig. 4).

The combined measurements of both mRNA and proteins from the same spatially sorted gates enabled a controlled comparison of protein and mRNA zonation patterns (Fig. 5b,c). Periportal bias (the difference between expression in the periporal and pericentral gates divided by mean expression) was significantly correlated between mRNA and proteins, indicating similar mRNA and protein zonation profiles for most genes (Spearman’s r = 0.39, P = 1.45 × 10^{-10} for all proteins; r = 0.45, P = 1.71 × 10^{-7} for highly expressed proteins; Fig. 5b,c). Notably, some genes exhibited discordant zonation of mRNA and proteins. These included genes that were zonated at the protein but not mRNA level, such as Rbp4, Idh3b, Mrpl43 and genes that were zonated at the mRNA but not at the protein level, such as A1cf, Clmm and Lsr (Fig. 5c). The discordant genes also included Hnf4a, a key hepatocyte transcription factor. The mRNA levels of Hnf4a were not zonated, whereas the protein content was higher in the periporal gates (Supplementary Fig 4e,f). This periporal...
protein bias is in line with previously reported involvement of HNF4A in periportal repression of Wnt-regulated pericentral genes\textsuperscript{17-19} and induction of periportally expressed targets\textsuperscript{20}. Indeed, we found that in Hnf4a knockout mice\textsuperscript{21}, periportal genes were preferentially downregulated compared to wild-type controls, whereas pericentral genes did not have preferential differential regulation (Fisher’s exact test $P = 1.16 \times 10^{-4}$; Supplementary Table 5). Thus, our analysis indicates that the majority of proteins and mRNA are similarly zonated, and highlight genes with potential post-transcriptional regulation.

**Zonation of hepatocyte miRNA content.** We next asked whether spatial sorting could be used to explore the regulatory mechanisms that shape hepatocyte zonation. miRNAs are short RNA
oligonucleotides, roughly 22 base pairs (bp) long, which target specific miRNAs through Watson–Crick base pairing, leading to increased degradation or decreased translation of target transcripts. Regulation by miRNAs seems to be important in liver development, metabolism and homeostasis. Notably, miRNA regulation may impact liver zonation, since mice lacking the miRNA central processing element Dicer in hepatocytes exhibit profound changes in zonation patterns. We reasoned that combined global measurements of the zonation profiles of miRNAs and mRNAs could identify potential miRNA-target regulatory interactions through the detection of miRNA-target pairs with anti-correlated expression profiles.

To this end, we performed miRNA microarray measurements of spatially sorted hepatocytes from three mice. We detected 302 miRNAs that were expressed in hepatocytes in all three mice. We further focused on 137 miRNAs that were classified as ‘high confidence’ in miRBase (v.22, downloaded 30 October 2018). Forty-five per cent (61 out of 137) of these high-confidence hepatocyte-expressed miRNAs were significantly zonated (FDR ≤ 0.2, Kruskal–Wallis test with Benjamini–Hochberg correction; Fig. 6a). Most zonation profiles (48 out of 61) were mildly pericentral with a centre of mass (COM) between 4 and 4.5, while 7 others showed strong periportal zonation (COM ≥ 6). We measured the expression of six of the miRNAs predicted to be zonated using quantitative PCR with reverse transcription (qRT–PCR), obtaining excellent correspondence with the microarray measurements (mean Spearman’s \( r = 0.70 \pm 0.24 \), Fisher’s method for combining \( P \) values of independent tests, \( P = 1.0 \times 10^{-15} \); Fig. 6b).

The zonated miRNAs included miRNAs previously described as playing a role in liver development, metabolism and regeneration. In agreement with previous studies, miR-122-5p, the most abundant miRNA in our measurements, comprised 46.5 ± 3.5% of the total miRNA content in hepatocytes. We found that miR-122-5p was peripherally zonated, with a 1.15-fold higher expression in the periportal miRNA content in hepatocytes. We found that miR-122-5p was significantly anti-correlated with its targets compared to randomized genes (Supplementary Table 7), indicating a potential regulatory role in shaping their zonation. Prominent pericentral miR-122-5p targets (genes that were repressed in their expression in the periportal layers where miR-122 was more abundant) included the canonical miR-122-5p target gene Csf, encoding citrate synthase, as well as Klf6 and Slc35a4 (ref. 29; Supplementary Fig. 5). MiR-30a-5p exhibited periportal zonation (periportal to pericentral ratio of 1.19, Kruskal–Wallis \( P = 0.007 \); Fig. 6a and Supplementary Table 6). Mtdh, a known target of miR-30a-5p, previously shown to change in expression in liver tumours, was pericentral and inversely zonated to its miRNA regulator (Spearman’s \( r = -0.81, P = 0.022 \); Supplementary Table 7). Additional zonated miRNAs included the pericentral miR-103-3p and miR-107-3p and the periporal miR-802-5p, which have been previously shown to modulate hepatic glucose sensitivity. In summary, our measurements revealed profound zonation of key hepatic miRNAs.

**Detection of putative miRNA-regulated hepatocyte target genes using zonation profiles.** The spatially stratified measurements of miRNAs and mRNAs could be used to identify potential miRNA regulation at the mRNA degradation level. Such regulation would be manifested in inverse correlations between the zonation profiles of a target mRNA and its regulating miR(s). To identify such interactions, we constructed a miRNA–mRNA regulatory network based on predictions from TargetScan (v.7.2) (Supplementary Table 7). We included all highly expressed hepatocyte genes and interactions with high confidence, and filtered out genes that are not predicted to be regulated by highly abundant miRNAs. The resulting network included 3,502 interactions between 131 miRNAs and 588 genes. For each gene, we constructed the cumulative regulating miRNA profile by summing up the zonation profiles of all miRNAs with a predicted regulatory interaction for the considered target gene (Supplementary Table 8). We computed the Spearman correlation between the gene’s mRNA zonation profile and the cumulative miRNA zonation profile and compared it to randomized degree-preserving networks (Fig. 7a).

Our analysis identified 11 genes that were significantly more anti-correlated with their regulating miRNAs compared to random (FDR < 0.2; Fig. 7b, Supplementary Table 8 and Supplementary Fig. 6). Pericentral target miRNAs included Acat1, which encodes the mitochondrially encoded subunit of mitochondrial complex IV (ref. 30), and R-spondin morphogens are secreted by pericentral liver endothelial cells, resulting in higher pericentral expression of Wnt-activated genes and lower pericentral expression of Wnt-inhibited genes. Notably, hepatocyte-specific Dicer knockout mice have perturbed zonation of Wnt-regulated genes, such as Glut1 and Arg1 (ref. 31). This suggests that miRNAs could differentially modulate hepatocyte Wnt signalling in different lobule zones. To explore this hypothesis, we analysed the miRNA-target subnetwork that includes genes associated with Wnt signal processing (see Methods and Fig. 7c,d). This analysis uncovered several key components of the Wnt network that exhibit zonation in hepatocytes and that have spatially anti-correlated zonated miRNAs. The Wnt receptors Fzd7 and Fzd8 were more highly expressed in pericentral hepatocytes, whereas their regulating miRNAs miR-149-5p, miR-30a-5p, miR-30a-3p, miR-21a-5p, miR-99a-5p and miR-100-5p were more abundant in periportal hepatocytes. In contrast, inhibitory components of Wnt signalling, such as Ctnmbp1 and Tcf7l1, were peripherally zonated. Tcf7l1, also known as Tcf3, is a transcriptional repressor of Wnt-activated genes that is inactivated by binding to \( \beta \)-catenin. This peripherally zonated gene is anti-correlated with its regulators miR-212-3p, miR-423-5p and miR-5107-5p (Fig. 7d). Ctnmbp1, which encodes \( \beta \)-catenin-interacting protein 1, prevents the binding of \( \beta \)-catenin to TCF7L1 and thus its removal and activation of Wnt target genes. The miR regulators of this periporal gene, miR-188-5p and miR-3102-5p, were pericentrally zonated (Fig. 7d). Thus, our analysis highlights zonated components of hepatocyte Wnt signalling and their potential regulation by miRNAs.

**Regulation of Wnt signalling components by miRNA.** Wnt is a major factor that shapes hepatocyte zonation. Wnt and R-spondin morphogens are secreted by pericentral liver endothelial cells, resulting in higher pericentral expression of Wnt-activated genes and lower pericentral expression of Wnt-inhibited genes. Notably, hepatocyte-specific Dicer knockout mice have perturbed zonation of Wnt-regulated genes, such as Glut1 and Arg1 (ref. 31). This suggests that miRNAs could differentially modulate hepatocyte Wnt signalling in different lobule zones. To explore this hypothesis, we analysed the miRNA-target subnetwork that includes genes associated with Wnt signal processing (see Methods and Fig. 7c,d). This analysis uncovered several key components of the Wnt network that exhibit zonation in hepatocytes and that have spatially anti-correlated zonated miRNAs. The Wnt receptors Fzd7 and Fzd8 were more highly expressed in pericentral hepatocytes, whereas their regulating miRNAs miR-149-5p, miR-30a-5p, miR-30a-3p, miR-21a-5p, miR-99a-5p and miR-100-5p were more abundant in periportal hepatocytes. In contrast, inhibitory components of Wnt signalling, such as Ctnmbp1 and Tcf7l1, were peripherally zonated. Tcf7l1, also known as Tcf3, is a transcriptional repressor of Wnt-activated genes that is inactivated by binding to \( \beta \)-catenin. This peripherally zonated gene is anti-correlated with its regulators miR-212-3p, miR-423-5p and miR-5107-5p (Fig. 7d). Ctnmbp1, which encodes \( \beta \)-catenin-interacting protein 1, prevents the binding of \( \beta \)-catenin to TCF7L1 and thus its removal and activation of Wnt target genes. The miR regulators of this periporal gene, miR-188-5p and miR-3102-5p, were pericentrally zonated (Fig. 7d). Thus, our analysis highlights zonated components of hepatocyte Wnt signalling and their potential regulation by miRNAs.

**MiRNA zonation and diseased liver states.** Many liver pathologies exhibit zonated damage, born out of the differential susceptibilities of periporal and pericentral hepatocytes to different insults. MiRNAs are attractive biomarkers due to their relative stability and high concentrations in the circulation. Therefore, we sought to explore whether the blood levels of zonated miRNA could be indicative of such zonated damage. Acetaminophen (APAP) intoxication leads to necrosis of pericentral hepatocytes. We analysed published miRNA levels measured in bulk liver and plasma of APAP-treated mice and found that almost all pericentral miRNAs were enriched in the plasma and depleted in the liver of APAP-treated mice.
Articles

**Fig. 6 | Zonated expression of hepatocyte miRNAs.** **a**, Mean expression versus zonation profile centre of mass (COM) for all detected high-confidence miRNAs. Selected miRNAs are labelled. The dashed red lines denote the median of each quantity. The red dots are miRNAs that are significantly zonated (two-sided Kruskal–Wallis test with Benjamini–Hochberg correction, $P < 0.2$). $n = 3$ mice. **b**, Validation of hepatocyte miRNA zonation profiles using qRT–PCR. The profiles for both qRT–PCR and microarrays are normalized by the expression levels of miR-103-3p ($r = 0.70 \pm 0.24$). The lines indicate the layer mean over three microarray mouse repeats (red) and the layer mean over three qRT–PCR mouse repeats (blue). The error bars indicate the s.e.m. The discrepancies between qRT–PCR and microarray profiles for let-7e-5p, miR-376a-3p and miR-802-5p may be due to the limited sensitivity of the microarray at low expression levels. $n = 3$ mice for the microarray and qRT–PCR.

Hepatocellular carcinoma is often associated with the activation of the Wnt pathway in hepatocytes. Wnt activity is pericentrally zonated in the healthy liver. Our study revealed zonated Wnt components, such as the pericentrally zonated Wnt receptors Fzd7 and...
**Fig. 7 | Network analysis of miRNA-target interactions.** a, Schematic illustration of the algorithm used to infer significant interactions between miRNAs and target genes. b, Zonation profiles of selected genes and their significantly anti-correlated cumulative miRNA profiles. The lines represent the mean of five mice for mRNA (green) and three mice for miRNA (yellow). The error bars represent the s.e.m. c, Regulatory network of hepatocyte-expressed Wnt pathway components and their expressed regulating miRNAs. Edges are coloured according to the correlation between miRNA and target. Edge weight is proportional to the absolute correlation value. n = 3 mice for the mRNA dataset and n = 5 mice for miRNA dataset. d, Selected pairs of miRNAs and their respective miRNA-target interactions. The transcripts of Ctnnbip1, Fzd8, Tcf7l1 and Znrf3 are anti-correlated with most of their regulating miRNAs, suggesting that miRNAs have a relatively more important role in regulating the expression of these genes in comparison to other genes. The lines represent the means of five mice for mRNA (green) and three mice for miRNA (yellow). The error bars represent the s.e.m.

**Fzd8** and the peripartally zonated Wnt inhibitors Ctnnbip1 and Tcf7l1. We found that these components are potentially modulated by zonated miRNA; thus, we hypothesized that perturbations in these miRNAs might be implicated in the carcinogenic process. To this end, we analysed the hepatocellular carcinoma dataset of the Cancer Genome Atlas (TCGA; Supplementary Table 9)\(^4\). We found...
that peripherally zonated miRNAs were downregulated whereas pericentrally zonated miRNA were upregulated (Spearman’s correlation $r = −0.29$ between the miRNA zonation profile COM and the log(fold change) between tumours and normal tissue, $P=0.0061$; Supplementary Fig. 8). For example, miR-99a and miR-100, which potentially inhibit Fzd8 expression in the periportal zones, were downregulated in tumours (log(fold change) = $−1.34$ and $−0.63$, respectively; Kruskal–Wallis $P=2.2 \times 10^{−16}$ and $P=2.6 \times 10^{−4}$, respectively). Conversely, mir-93, which potentially inhibits the Wnt inhibitor Tcf711 in the pericentral zone, was upregulated (log(fold change) = 1.49, $P=5.4 \times 10^{−26}$; Supplementary Fig. 8).

The anti-correlation between miRNA fold change in hepatocellular carcinoma and its zonation in the healthy liver became even stronger when focusing on patients without Wnt-activating somatic mutations (Spearman’s $r = −0.36$, $P=4.5 \times 10^{−3}$; see Methods). Such increased inverse correlation could indicate that miRNAs may play a role in activating the Wnt pathway independently of somatic mutations in the Wnt pathway.

**Discussion**

The liver exhibits profound division of labour among hepatocytes that reside at different zones. Thus, understanding and modelling liver function requires characterization of hepatocyte functions at each lobule coordinate. In this study, we present spatial sorting, a generic approach used to isolate large amounts of hepatocytes with high spatial resolution for a broad range of downstream measurement modalities. The approach employs zonated surface markers that can be identified by spatially resolved transcriptomic atlases. We demonstrate applications of this approach for resolving the zonation of hepatocyte proteins and miRNAs. The approach can be readily applied to other structured organs and cells types exhibiting zonation, including liver endothelial cells, intestinal enterocytes and kidney cells. The use of endogenous surface markers renders spatial sorting particularly useful for studying zonation in humans.

Our proteome analysis revealed some notable discordance between the average hepatocyte levels of proteins and miRNAs, mostly for genes encoding secreted proteins (Fig. 4). In contrast, between the average hepatocyte levels of proteins and mRNAs, spatial sorting particularly useful for studying zonation in humans.

We identified pericentral zonation of the key Wnt receptors Fzd7 and Fzd8 and perportal zonation of the Wnt inhibitors Tcf711 and Ctnnb1. This joins previous reports of periportal zonation of APC, a key Wnt regulator. Our study further identified miRNAs that regulate these zonated Wnt components. Thus, miRNAs seem to be upstream of Wnt signalling. These results could explain the effects of hepatocyte-specific Dicer and β-catenin knockout. While Dicer knockout resulted in perturbed zonation of Wnt targets, β-catenin knockout did not substantially alter miRNA levels.

Our finding of decreased expression of miRNAs that may potentially inhibit Wnt receptors (mir-99a and mir-100) in hepatocellular carcinoma could indicate that Wnt-targeting miRNAs may play a role in the liver carcinogenic process.

Our approach allowed us to obtain up to a few hundreds of thousands of hepatocytes per sorted population. While this amount is compatible with a broad range of assays, it is insufficient for assays that require massively larger amounts of material, such as RNA methylation and metabolic profiling. Moreover, since the approach is FACS-based, measuring metabolites, which are labile, would be compromised by the substantial incubation periods involved in the protocol. Nevertheless, it will be interesting to apply spatial sorting to explore additional zonated hepatocyte features, including chromatin modifications, DNA methylation, three-dimensional chromosomal conformations, DNA mutation spectra and chromosomal aberration. Such measurements could resolve hepatocyte cell identity, regulatory mechanisms and susceptibility to damage in each zone.

**Methods**

**Animal experiments.** Mouse experiments were approved by the Institutional Animal Care and Use Committee of the Weizmann Institute of Science and performed in accordance with institutional guidelines. The sorting experiments were conducted on five 3-month-old C57BL/6JolaHsd male mice obtained from Envigo. Mice were kept in a reverse light–dark cycle. Mice were anaesthetized with ketamine (100mg kg$^{−1}$) and xylazine (10mg kg$^{−1}$) dissolved in 1x PBS and injected intraperitoneally 6–9h after lights off (Zeitgeber time 18–21).

For the imaging experiments, the livers of the 3-month-old C57BL/6JolaHsd male mice were collected and fixed in cold parafomaldehyde for 3h at 4°C overnight. Antibodies used were Alexa Fluor 647 rat anti-mouse CD73 (catalogue no. 561543; BD Biosciences) and Alexa Fluor 555 mouse anti-E-cadherin (catalogue no. 560064; BD Biosciences). On the next day, slices were washed with

---

NATURE METABOLISM | VOL 1 | SEPTEMBER 2019 | 899–911 | www.nature.com/natmetab

---
Articles

PBST three times. Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI; 1:100 in PBS, 10 min). Imaging of the liver portal-central axis was performed on a Nikon Ti-E inverted fluorescence microscope with a 100× oil-immersion objective and a Puxis 1024 charge-coupled device camera (Photometrics) using the MetaMorph software (v.7.7.11.0) in scan stage option.

The Z-projected images of lobule scans (8 scans, 3 mice) were analysed.

Membrane segments of hepatocytes were measured for the intensity of Alexa Fluor 555 (E-cadherin) and Alexa Fluor 647 (CD73). Background, set as the paired cytoplasmic intensity for each membrane signal, was subtracted. Segments were then binned into eight groups representing eight lobule layers (1 marks the pericentral-most layer, and 8 marks the perportal-most layer), according to their radial distance from the central vein. The median intensity of the segments from each lobule layer was calculated and averaged over the different lobules (Fig. 3b,c). Values were scaled from 0 to 1 and the plot was smoothed with a sliding window of 3.

Immunofluorescence of aldolase B and glutamine synthetase.

Immunofluorescence of aldolase B (catalogue no. PA5-30218; Thermo Fisher Scientific) and glutamine synthetase (catalogue no. MA5-27749; Thermo Fisher Scientific) was performed in a similar fashion to E-cadherin and CD73; however, secondary antibodies were needed after overnight incubation with primary antibodies. Slides were incubated for 2 h at room temperature with secondary antibodies before staining with DAPI. For aldolase B, we used 1:500 goat anti- rabbit immunoglobulin G (IgG) (H+L) crossadsorbed secondary antibody conjugated to cyanine 5 (catalogue no. A10520; Thermo Fisher Scientific).

Immunofluorescence of HNF4A.

HNF4A antibody staining was done according to a previously published protocol35, Paraffin-embedded liver blocks were sectioned into 10 μm thick slices. Slides were deparaffinized and rehydrated, followed by methanol incubation for 5 min. Slides were then incubated with PBS, 10% BSA and 0.3% Triton-X100, after which slides were incubated overnight with HNF4A antibody (catalogue no. PP-H1415-00; R&D Systems) in a dilution of 1:350 in blocking solution at 4 °C. Slides were incubated for 2 h in 1:700 donkey anti-Mouse IgG (H+L) highly crossadsorbed secondary antibody conjugated to Alexa Fluor 594 (catalogue no. A-21203; Thermo Fisher Scientific) with 1:1,000 DAPI.

Liver perfusions and hepatocytes dissociation.

Once mice had been anaesthetized, their livers were perfused as described previously36, with a few adjustments. A 27 G syringe connected to the perfusion line and pump was inserted into the vena cava; 25 ml of pre-warmed to 37 °C EGTA buffer followed by 25 ml of pre-warmed to 37 °C enzyme buffer solution (EBS) with 2.3 U of Liberase Blendzyme 3 recombinant collagenase (Roche) were cannulated into the vena cava. Shortly after the beginning of the perfusion, the portal vein was cut to allow drainage of the blood.

After perfusion of 10 ml EBS and 15 ml of enzyme solution, livers were explanted into a Petri dish with 25 ml of pre-warmed EBS and gently minced using forceps. Dissociated liver cells were collected and filtered through a 100 μm cell strainer. Cells were spun down at 300 × g for 3 min at 4 °C to obtain the hepatocyte-enriched pellet. Pellet was resuspended in 25 μl cold EBS.

Flow cytometry and cell sorting.

Cells were sorted into lysis buffer supplied with the Dynabeads mRNA DIRECT Micro Purification Kit (Invitrogen). RNA was extracted according to the protocol provided with the kit: 2 μl of the extracted mRNA from each sample was used for the libraries. Library preparation was done according to the mcSCRB-seq protocol35. The complementary DNA was pre-amplified with 10–15 cycles, depending on the cDNA concentration indicated by qRT–PCR quality control; 2 ng of the amplified cDNA was converted into the sequencing library with the Nextera XT DNA Library Preparation Kit (catalogue no. EC-131-124; Illumina), according to the protocol supplied. Quality control of the resulting libraries was performed with a High Sensitivity DNA ScreenTape Analysis system (catalogue no. 5067-5584; Agilent Technologies). Libraries were loaded with a concentration of 2.2 pmol on 75-cycle high-output flow cells (catalogue no. FC-404-206-03) at NextSeq 500 (Illumina) with the following cycle distribution: 87 bp index1, 16 bp read1, 66 bp read2 (no index2 needed). A total of 40 libraries (8 sorted populations for 5 different mice) were sequenced.

Analysis of potential NPC representation. Although we used negative selection with surface markers for endothelial (CD31) and immune cells (CD45), we compensation was corrected manually. To collect eight populations, each enriched with spatially stratified hepatocytes with equal viability and ploidy levels, events were screened through the following five nested gates (Fig. 3a,b and Supplementary Fig. 1b–d): (1) hepatocyte gates from all events—set by plotting forward side scatter area (FSC-A) against side scatter area (SSC-A) and excluding large clusters and small debris; (2) singlets FSC—set by excluding the margins of FSC-A and FSC width plot; (3) singlets SSC—excluding the upper margins of SSC width when plotted against SSC-A; (4) live cell gates according to the Zombie Green–Alexa Fluor 488–propidium iodide (PI) double staining; and (5) hepatocytes (HNF4A+ and CD45+). HNF4A staining was done according to the protocol provided with the kit: 2 μl of the extracted mRNA from each sample was used for the libraries. Library preparation was done according to the mcSCRB-seq protocol35. The complementary DNA was pre-amplified with 10–15 cycles, depending on the cDNA concentration indicated by qRT–PCR quality control; 2 ng of the amplified cDNA was converted into the sequencing library with the Nextera XT DNA Library Preparation Kit (catalogue no. EC-131-124; Illumina), according to the protocol supplied. Quality control of the resulting libraries was performed with a High Sensitivity DNA ScreenTape Analysis system (catalogue no. 5067-5584; Agilent Technologies). Libraries were loaded with a concentration of 2.2 pmol on 75-cycle high-output flow cells (catalogue no. FC-404-206-03) at NextSeq 500 (Illumina) with the following cycle distribution: 87 bp index1, 16 bp read1, 66 bp read2 (no index2 needed). A total of 40 libraries (8 sorted populations for 5 different mice) were sequenced.

Sequencing analysis pipeline.

Illumina output files were demultiplexed with bcl2fastq v2.17.1, the resulting FASTQ files from the mRNA sequencing experiments were analysed with the DIABLO pipeline34. Using STAR (v.2.5.3a), reads were aligned to a transcriptome index of the GRCm38 (release 84; Ensembl) and exonic unique transcripts were extracted. Sequencing reads were then binned into eight groups representing eight lobule layers, reducing spatial accuracy (Supplementary Fig. 1c,d).

Mass spectrometry.

Forty samples (five mice, eight populations each) were digested by trypsin and analysed with liquid chromatography–tandem mass spectrometry on a Q Exactive Plus Hybrid Quadrupole–Orbitrap Mass Spectrometer (Thermo Fisher Scientific). The data were analysed with MaxQuant v.1.5.2.8 (ref. 40) against the mouse Uniprot database. Data were quantified using the same software. We retained proteins with a FDR <0.01 in at least 2 samples in one of the 8 groups, identified by at least 2 peptides across all samples; 3,210 protein groups were identified. For each sample, label–tag quantification intensities for each peptide were normalized by the sum of all intensities, yielding the expression fraction out of the total protein detected.

Analysis of potential NPC representation. Although we used negative selection with surface markers for endothelial (CD31) and immune cells (CD45), we
computationally validated that there was no significant presence of NPCcs in our sorted populations. To this end, we compiled a gene expression dataset of nine major liver cell types from previous publications11,12, including endodermal cells, T cells, plasmacytoid dendritic cells, Kupffer cells, liver capsule macrophages, B cells, neutrophils, hepatocytes13 and cholangiocytes14. We filtered the dataset to include only genes with a transcriptome fraction of at least $10^{-4}$ in at least one cell type, and which were present in our spatially sorted hepatocyte RNA-seq data, resulting in 9,805 expressed genes. We next identified the top 100 marker genes for each of the nine cell types, defined as having the largest ratio between their expression in a given cell type and the maximum expression over all other cell types (Supplementary Table 1). The summed expression of the marker genes, which we termed the cell types’ signature, was calculated for both the transcriptomic and proteomic data of our sorted populations. In the transcriptomic data, the hepatocyte signature made up 15.7±0.05 of the total protein content, while each of the other cell types’ signatures comprised only 0.004±0.4113% of the protein content. Likewise, there were no significant differences in the levels of the proteomic cell type signatures across the spatially sorted populations (Kruskal–Wallis $P=1$).

Calculating zonation of Hnf4a targets. We used the data from Holloway et al.21 for differentially expressed genes in male mice with Hnf4a knockout compared to wild-type male mice. We extracted a unique list of these genes and calculated the Benjamini–Hochberg FDR to identify significantly differentially expressed genes, set to have an FDR<0.01. Those genes were intersected with our spatial sorting RNA-seq data, resulting in 1,166 genes. We next grouped the genes according to the following categories: genes were considered downregulated or upregulated if their knockout/wild-type log2 fold change was smaller than −1 or greater than 1, respectively; genes were considered pericentral/periporal, according to our spatial sorting RNA-seq data, if their zonation COM was smaller or larger than 4.5, respectively. These criteria formed 4 groups: upregulated periporal genes (296 genes); upregulated portal periporal genes (123 genes); downregulated pericentral genes (348 genes); and downregulated portal periporal genes (245 genes). We used Fisher’s exact test to identify a non-random association between differential expression and zonation under Hnf4a knockout ($P=1.2 \times 10^{-5}$).

miRNA microarrays. Total RNA (100 ng), isolated from bulk populations of 50,000 spatially sorted hepatocytes ($n=3$) per FACS gate, was labelled with cyanine 3 during transformation into cDNA using an miRNA Labelling Kit (Agilent Technologies) and Spike Kit (Agilent Technologies). CDNA was hybridized to the Mouse miRNA Microarray, release 21.0, 8 x 60 K (v.21) microarray slides (Agilent Technologies) according to the Agilent miRNA Hybridization Kit protocol (Agilent Technologies) and scanned using a G2505B array scanner (Agilent Technologies). Data were extracted using the Feature Extraction Software (v.10.7.3.1, Agilent Technologies) with default parameters.

qRT–PCR. The total RNA isolated from bulk populations of 50,000 spatially sorted hepatocytes ($n=3$) per FACS gate was diluted to 5 ng/μl and cDNA was reverse-transcribed using the miRCURY LNA RT Kit (catalogue no. 339340; QIAGEN), according to the manufacturer’s instructions, on a ProFlex PCR System (Applied BioSystems). Plates were prepared using the miRCURY LNA SYBR Green PCR Kit (catalogue no. 339346; QIAGEN; Supplementary Table 10). Each 10 μl reaction volume contained 5 μl 2x miRCURY SYBR Green Master Mix, 0.5 μl 6-carboxy-X-rhodamine reference dye, 1 μl PCR primer mix, 0.5 μl RNase-free water and 3 μl of cDNA sample diluted 1:60. qRT–PCR reactions and measurements were performed on a StepOne Real-Time PCR System (catalogue no. 4373999; Applied Biosystems) according to the manufacturer’s instructions. Correlations between qRT–PCR and microarray measurements were calculated on a mouse-by-mouse basis.

Centre-of-mass calculation. The COM of an expression profile $x$ (spread over $z=1.8$ FACS gates) was calculated as:

$$\text{COM}(x) = \frac{\sum_{i=1}^{n} z_i x_i}{\sum_{i=1}^{n} z_i}$$  

This formula yields a COME[1,8] number that shows around which gate most of the expression is distributed.

Comparing bulk mRNA with published single-cell RNA-seq. Spatial sorting produces subpopulations of hepatocytes that are enriched for specific lobule layers; however, each FACS gate includes several lobule layers. Thus, to compare the bulk mRNA measurements of the FACS-gated subpopulations to the zonation measurements previously reconstructed using spatially resolved single-cell transcriptomics (Fig. 3c), we computationally estimated the relative abundances of the different lobule layers in each of the FACS gates. To this end, we implemented Cibersort (https://cibersort.stanford.edu/)23. We extracted a gene signature list for each layer from the single-cell RNA-seq data. A total of 17 genes with a mean expression $> 5 \times 10^{-3}$, a zonation FDR < 0.01 and dynamic range of at least tenfold between the mean of the two periporal layers and the mean of the two pericentral layers were used for the analysis. The means of five mice for the zonation profiles of the three layers across the eight FACS gates were used as the mixed populations dataset. The relative abundances of each of the nine layers in each of the eight sorted populations was calculated with the ‘disable quantile normalization’ option ticked. Figure 3c shows the mean expression in each FACS gate over five mice (blue) and the mean expression in single-cell RNA-seq data, weighted by the relative abundances of each layer in each FACS gate (yellow).

Comparing proteins and RNA. Out of the 3,210 proteins (Supplementary Table 3) detected using mass spectrometry and 14,027 mRNAs (Supplementary Table 2) detected using RNA-seq, 3,051 were found in both datasets (Supplementary Table 4). The median expression fraction of five mice was calculated for each gate in each measurement. We produced scatterplots showing the averages over all gates of the eight mRNA and eight protein medians for every gene and found a Spearman correlation of $r=0.50$ (0.48–0.50 for each gate independently). To better characterize mRNA and protein ratios in different KEGG pathways24, we plotted a regression line of protein by mRNA. The residual of the proteins from the regression line was calculated and grouped according to KEGG pathways (Fig. 4b and Supplementary Fig. 2b).

Computing zonation. For each of the 3,051 common proteins and mRNAs, a Kruskal–Wallis test was performed to check for variability between different sorted gates ($n=40$ populations, 5 mice in each of the 8 gates, d.f.=37). To correct for multiple hypotheses, we used the Benjamini–Hochberg correction to obtain the FDR for each hypothesis. We classified proteins as zonated if they had an FDR<0.05; 1,672 were significantly zonated. To produce the protein zonation heatmap (Fig. 5a), we first removed all proteins with a median of label-free quantification $\leq 2$ (ref. 25) in any of the eight gates (479 proteins). Next, we normalized all protein profiles to their maximum across all FACS gates and sorted them by their COM (Fig. 5a).

Microarray data. Only miRNAs that were annotated as ‘high confidence’ in mirBase v.22 (downloaded 30 October 2018) were kept for analysis. The raw signal for each miRNA in each FACS-gated and in each array was normalized by the total signal per gate per array. Only miRNAs present in all three biological replicates were retained for further analysis and their initial normalized signal was averaged over all arrays. Finally, the averaged signal was divided again by the total signal in each gate. (This operation amounted to dividing by a number very close to 1, since only miRNAs with very low expression were not present in some of the replicates.) MiRNA zonation was inferred using the Kruskal–Wallis test (for each miR, comparison of 8 gates with each having 3 replicates, d.f.=21), and applying a Benjamini–Hochberg correction on the $P$ values obtained from the Kruskal–Wallis test. MiRNAs with an FDR$=0.2$ were classified as zonated.

Differential zonation of miR-122-5p targets. The targets of miR-122-5p were taken from Tsai et al.14; 146 of the targets listed were expressed in our liver zonated transcriptome data. The mean COM was calculated for these targets and for 1,000 random samplings (with replacement) of 146 liver-expressed genes (genes with maximal expression over all gates that was at least $4 \times 10^{-4}$ of the transcriptome). We performed a Wilcoxon rank-sum test for the COMs of the target genes versus the COMs of randomly sampled genes, yielding $P=0.018$.

Analysis of APAP-induced liver injury data. Mature sequence suffixes (-3′/5′-5′) were removed from the miRNA names in the dataset from Wang et al. and names were matched with our miRNA zonation data, after merging miRNAs with identical names. MirBase names were defined as those with a maximal expression over all gates that was at least 4 and Kruskal–Wallis FDR $\leq 0.2$, whereas periporal miRNAs had a COM $> 4.5$ and Kruskal–Wallis FDR $\leq 0.2$. A Wilcoxon’s rank-sum test was calculated for the ratios log(treatment/control) of all detected pericentral miRNAs in the plasma versus in the liver.

Analysis of TCGA hepatocellular carcinoma data. TCGA hepatocellular carcinoma data (level 3 mRNA and miRNA-seq datasets, not RSEM-normalized) were downloaded using the R package https://rtcgat.github.io/RTCGA/, v.1.5.1. To match our microarray miRNA nomenclature, which can inform from which mature miRNA the signal came from (that is, the -3′/5′ suffixes to the miRNA names), and the RNA-seq data, which do not have this specificity, names and expression values were merged between degenerate miRNA names. Specifically, -3′/5′ suffixes and genomic locus variants (indicated by additional digits after the canonical miRNA number) were removed from the miRNA names and the expression values of miRNAs with the resulting identical name were summed. The resulting TCGA data included the values of 1,046 miRNAs over...
372 tumour samples and 50 normal samples. For each miRNA, we computed the log, ratio of the median expression levels in the tumour and normal samples and examined the Spearman correlations of the log, ratios and the zonation profile. COM for the miRNAs that were included in our spatial sorting study. To examine the magnitude of fold change in miRNA expression independently of Wnt pathway mutations, we stratified the TCGA data by patients’ mutation state and removed from the analysis all patients with miRNAs in at least one of the genes APC and CTNNB1, and all FZD genes (FZD1–FZD10).

MIrNA-target network construction, randomized networks and genes–miRNA anti-correlation. All miRNA-target interaction predictions were downloaded from TargetScanMouse v.7.2 (ref. 3). (data released in August 2018, downloaded 24 October 2018), including conserved and non-conserved sites. Predicted edges were filtered for only liver-expressed miRs, annotated ‘high confidence’ in miBase v.22, weighted context++ score percentile ≥95 and genes expressed with maximum fraction (over genes) of total transcriptome ≥4×10^−5. The resulting network included 33,672 interactions between 131 miRNAs and 6,650 genes. For each gene g we constructed the cumulative expression profile M_g(z) of all miRNAs p,∀i∈{1,2,…,N_g},z=1,…,8) predicted to target it:

\[
M_g(z) = \sum_{i=1}^{N_g} \mu_i(z)
\]

We calculated the Spearman correlation between the expression profile x(z) of each gene g and M_g(z). For the network randomization procedure, we only included highly expressed genes (maximal expression ≥10^3) and genes that are predicted to be regulated by a substantial fraction of the miRNA transcriptome (mean expression of the cumulative miRNA profile ≥10^3). This network included 588 genes regulated by 131 miRNAs with 3,502 interactions. We then created 1,000 networks with randomized edge assignment using mfinder (v.1.2)73 with examined the Spearman correlations of the log2 ratios and the zonation profile included highly expressed genes (maximal expression

The liquid chromatography–tandem mass spectrometry proteomic data were uploaded to the ProteomeXchange via the PRIDE archive, with project identifier PXD014512. Processed data can be found in Supplementary Table 5. The miRNA microarray data have been deposited with the NCBI Gene Expression Omnibus under accession code GSE134827. Supplementary Table 6 summarizes the miRNA data.

Received: 20 January 2019; Accepted: 8 August 2019; Published online: 16 September 2019

References

1. Hoehme, S. et al. Prediction and validation of cell alignment along microvessels as order principle to restore tissue architecture in liver regeneration. Proc. Natl Acad. Sci. USA 107, 10371–10376 (2010).
2. Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing diploid Axin2+ cells fuel homeostatic renewal of the liver. Nature 524, 180–185 (2015).
3. Colnot, S. & Perret, C. in Molecular Pathology of Liver Diseases (ed. Monga, P. S.) 7–16 (Springer, 2011).
4. Ben-Moshe, S. & Izikovitz, S. Spatial heterogeneity in the mammalian liver. Nat. Rev. Gastroenterol. Hepatol. 16, 395–410 (2019).
5. Gebhardt, R. Metabolic zonation of the liver: regulation and implications for liver function. Pharmacol. Ther. 53, 275–354 (1992).
6. Jungernann, K. & Keitzmann, T. Zonation of parenchymal and nonparenchymal metabolism in liver. Annu. Rev. Nutr. 16, 179–203 (1996).
7. Halpern, K. R. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 542, 352–356 (2017).
8. Lindros, K. O. & Penttila, K. W. Digitonin-collagenase perfusion for efficient separation of periporal or perivenous hepatocytes. Biochem. J. 228, 757–760 (1985).
9. Quisbroff, C., Grunnet, N. & Cornell, N. W. Digitonin perfusion of rat liver. A new approach in the study of intra-acinar and intracellular compartmentation in the liver. Biochem. J. 226, 289–297 (1985).
10. Budnik, B., Levy, E., Harmange, G. & Slavov, N. ScCoPE: MS mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. Genome Biol. 19, 161 (2018).
11. Buenrostro, J. D. et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490 (2015).
12. Doi, Y. et al. Development of complementary expression patterns of Es- and N-cadherin in the mouse liver. Hepatol. Res. 37, 230–237 (2007).
13. Halpern, K. R. et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. Nat. Biotechnol. 36, 962–970 (2018).
14. Azimifar, S. B., Nagaraj, N., Cox, J. & Mann, M. Cell-type-resolved quantitative proteomics of murine liver. Cell Metab. 20, 1076–1087 (2014).
15. Schwanhäuser, B. et al. Global quantification of mammalian gene expression control. Nature 473, 337–342 (2011).
16. Odom, D. T. et al. Control of pancreas and liver gene expression by HNF transcription factors. Science 303, 1378–1381 (2004).
17. Torre, C., Perret, C. & Colnot, S. Molecular determinants of liver zonation. Prog. Mol. Biol. Transl. Sci. 97, 127–150 (2007).
18. Stanulović, V. S. et al. Hepatic HNF4α deficiency induces periportal expression of glutamine synthetase and other pericentral enzymes. Hepatology 45, 434–444 (2007).
19. Colletti, M. et al. Convergence of Wnt signaling on the HNF4α-driven transcription in controlling liver zonation. Gastroenterology 137, 660–672 (2009).
20. Brosch, M. et al. Epigenomic map of human liver reveals principles of zonated morphogenic and metabolic control. Nat. Commun. 9, 4130 (2018).
21. Holloway, M. G., Miles, G. D., Dombkowski, A. A. & Waxman, D. J. Liver-specific hepatocyte nuclear factor-4α deficiency: greater impact on gene expression in male than in female mouse liver. Mol. Endocrinol. 22, 1274–1286 (2008).
22. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466, 835–840 (2010).
23. Lagos-Quintana, M. et al. Identification of tissue-specific microRNAs from mouse. Curr. Biol. 12, 735–739 (2002).
24. Kota, J. et al. Therapeutic microRNA delivery suppresses tumorigenesis in a mouse liver cancer model. Cell 137, 1005–1017 (2009).
25. Sekeina, S., Ogawa, R., Mcmanus, M. T., Kanai, Y. & Hebrok, M. Dicer is required for proper liver zonation. Nat. Commun. 6, 4130 (2015).
26. Sekine, S., Ogawa, R., Mcmanus, M. T., Kanai, Y. & Hebrok, M. Dicer is required for proper liver zonation. J. Pathol. 219, 365–372 (2009).
27. Arvey, A., Larsson, E., Sander, C., Leslie, C. S. & Marks, D. S. Target mRNA abundance dilutes microRNA and siRNA activity. Mol. Syst. Biol. 6, 363 (2010).
28. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42, D68–D73 (2014).
29. Bissell, U. et al. Absolute quantification of microRNAs by using a universal reference. RNA 15, 2375–2384 (2009).
30. Tsi, W.-C. et al. MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis. J. Clin. Invest. 123, 2884–2897 (2012).
31. Li, W. E., Dai, H., Ou, Q., Zuo, G.-Q. & Liu, C. A. Overexpression of microRNA-30a-5p inhibits liver cancer cell proliferation and induces apoptosis by targeting MTDH/PTEN/AKT pathway. Tumour Biol. 37, 5885–5895 (2016).
32. Kornfeld, J.-W. et al. Obesity-induced overexpression of miR-802 impairs glucose metabolism through silencing of Hif1β. Nature 494, 111–115 (2013).
33. Trajkovski, M. et al. MicroRNAs 103 and 107 regulate insulin sensitivity. Nature 474, 649–653 (2011).
33. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4, e00905 (2015).

34. Benhamouche, S. et al. Apc tumor suppressor gene is the “zoonation-keeper” of mouse liver. *Dev. Cell* 10, 759–770 (2006).

35. Birchmeier, W. Orchestrating Wnt signalling for metabolic liver zonation. *Nat. Cell Biol.* 18, 463–465 (2016).

36. Burke, Z. D. & Tosh, D. The Wnt/β-catenin pathway: master regulator of liver zonation? *Bioessays* 28, 1072–1077 (2006).

37. Gebhardt, R. & Hovhannisyan, A. Organ patterning in the adult stage: the role of Wnt/β-catenin signaling in liver zonation and beyond. *Dev. Dyn.* 239, 45–50 (2010).

38. Thompson, M. D. & Monga, S. P. WNT/β-catenin signaling in liver health and disease. *Hepatology* 45, 1298–1305 (2007).

39. Yang, J. et al. β-catenin signaling in murine liver zonation and regeneration: a search for the Wnt-β-catenin-Hepatocyte growth factor axis. *Hepatology* 2, 845–860 (2018).

40. Planas-Paz, L. et al. The RSPO-LGR4/5-ZNRF3/RNF43 module controls liver zonation and size. *Nat. Cell Biol.* 18, 467–479 (2016).

41. Preziosi, M., Okabe, H., Poddar, M., Singh, M. & Monga, S. P. Endothelial Wnts regulate β-catenin signaling in murine liver zonation and regeneration: a search for the Wnt-β-catenin-Hepatocyte growth factor axis. *Hepatology* 2, 845–860 (2018).

42. Rocha, A. S. et al. The angiocrine factor Rspo3in is a key determinant of liver zonation. *Cell Rep.* 13, 1757–1764 (2015).

43. Shy, B. R. et al. Regulation of Tcf771 DNA binding and protein stability as principal mechanisms of Wnt/β-catenin signaling. *Cell Rep.* 4, 1–9 (2013).

44. Iago, K. et al. Inhibition of Wnt signaling by ICAT, a novel β-catenin-interacting protein. *Genes Dev.* 14, 1741–1749 (2000).

45. Wang, K. et al. Circulating microRNAs, potential biomarkers for drug-induced liver injury. *Proc. Natl Acad. Sci. USA* 106, 4402–4407 (2009).

46. Zucman-Rossi, J., Villanueva, A., Nault, J.-C. & Llovet, J. M. Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology* 149, 1236–1239.e1 (2015).

47. Chu, A. et al. Large-scale profiling of microRNAs for The Cancer Genome Atlas. *Nucleic Acids Res.* 44, e3 (2016).

48. Moor, A. E. et al. Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell* 175, 1156–1167.e15 (2018).

49. Park, J. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* 360, 758–763 (2018).

50. Guder, W. G. & Ross, B. D. Enzyme distribution along the nephron.

51. Xu, H. et al. Liver-enriched transcription factors regulate microRNA-122 that is induced 24 hours after liver injury. *J. Hepatol.* 64, 1051–1058 (2016).

52. Roderburg, C. et al. Micro-RNA profiling reveals a role for miR-29 in human induced liver injury. *Proc. Natl Acad. Sci. USA* 105, 10513–10518 (2008).

53. Farid, W. R. et al. Hepatocyte-derived microRNAs as serum biomarkers of hepatic injury and rejection after liver transplantation. *Liver Transpl.* 18, 290–297 (2012).

54. Bagnoli, J. W. et al. Sensitive and powerful single-cell RNA sequencing using mSCRR-seq. *Nat. Commun.* 9, 2397 (2018).

55. Holzhütter, H.-G., Drasdo, D., Preusser, T., Lippert, J. & Henney, A. M. The principal mechanisms of Wnt/β-catenin signaling and their use in investigating mechanisms of hepatotoxicity, cell signaling and therapeutic modalities. *Cell Rep.* 28, 239–249 (2020).

56. Robinson, M. D. & Oshlack, A. A scaling normalization method for RNA-seq data. *Genome Biol.* 11, R25 (2010).

57. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010).

58. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372 (2008).

59. Li, B. et al. Adult mouse liver contains two distinct populations of cholangiocytes. *Cell Stem Cell* 9, 478–489 (2017).

60. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457 (2015).

61. Kusnetsov, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30 (2000).

62. Mederacke, I., Dapito, D. H., Affò, S., Uchinami, H. & Schwabe, R. F . Molecular clocks: virtual liver: a multidisciplinary, multilevel challenge for systems biology. *Cell Tissue Res.* 368, 261–270 (2017).

63. Liebermeister, W. et al. Visual account of protein investment in cellular functions. *Proc. Natl Acad. Sci. USA* 111, 8488–8493 (2014).

64. Otto, A. et al. Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat. Commun.* 1, 137 (2010).

Acknowledgements

We thank K. Haga and the Flow Cytometry Unit (Weizmann Institute of Science) for FACS technical support; T. Ziv and the Smoer Proteomics Center (Technion) for performing the liquid chromatography–tandem mass spectrometry and analysing the results; and D. Pilzer and the Genomic Technologies Unit (Weizmann Institute of Science) for performing the miRNA microarray measurements. We thank all members of the laboratory for valuable comments. S.I. is supported by the Henry Chanoch Krenter Institute for Biomedical Imaging and Genomics, The Leir Charitable Foundations, the Richard Jakabshind Laboratory of Systems Biology, the Cymerman Jakabshind Prize, The Lord Sieff of Brimpton Memorial Fund, the I-CORE programme of the Planning and Budgeting Committee (grant no. 1902/12), the Israel Science Foundation (grant no. 1486/16), the Broad Institute–Israel Science Foundation (grant no. 2615/18), the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant no. 768956), the Bert L. and N. Kugie Valley Foundation, the Howard Hughes Medical Institute (HHMI) International Research Scholar Award (grant no. 5508734), The Wolfson Family Charitable Trust (grant no. 21376) and the Edmond de Rothschild Foundations.

Author contributions

K.B.H. and S.I. conceived the study. S.R.M. and S.I. designed the experiments. S.R.M. prepared all the samples. S.R.M. and Y.S. analysed the data. A.E.M. contributed to the data analysis. R.M. and T.V. assisted with the immunofluorescence experiments. K.B.H. contributed to establishing the methodology. S.I. supervised the study. S.R.M., Y.S. and S.I. wrote the manuscript. All authors reviewed the manuscript and provided input.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41225-019-0109-9. Correspondence and requests for materials should be addressed to S.I. Peer review information Primary Handling Editor: Elena Bellafante. Reprints and permissions information is available at www.nature.com/reprints. © The Author(s), under exclusive licence to Springer Nature Limited 2019
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's $d$, Pearson's $r$), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection
For RNA seq analysis, we used BCL2fastq tool (Illumina), STAR (2.5.3a) for alignment and zUMIs for quantification.
For proteomics analysis, we used MaxQuant (1.6.0.16) for identification and quantification.
For miRNA microarrays analysis, we used Agilent Feature Extraction software.
we used BD FACSDiva™ Software (BD Biosciences-US) for sorting and MetaMorph software for microscope image acquisition.

Data analysis
Data analysis Data were analyzed using MATLAB (R2018) and RStudio (v1.2, using R v3.5). R packages used for analysis are "dplyr",tidyr","biomaRt" and "edgeR". Additional tools employed are mfinder, CIBERSORT and Proteomap.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA sequencing data that support the findings of this study have been deposited in NCBI Sequence Read Archive (SRA) with the BioProject accession code PRJNA565572, and the SRA identifiers SAMN12360372 - SAMN12360382. Supplementary Table 2 summarizes the UMI counts per million for each sample. Supplementary Table 11 summarizes the zUMI barcodes used per each sample and the corresponding zUMI settings.
LC-MS/MS proteomic data was uploaded to ProteomeXchange via the PRIDE database, with the project identifier PXD014512. Processed data can be found in Supplementary Table 3.
MiRNA microarray data have been deposited in NCBI Gene Expression Omnibus (GEO) with the primary accession code GSE134827. Supplementary Table 6
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size: n=5 mice in RNA-seq and MS/MS experiments. Sample size of 5 was chosen to allow power for non-parametric statistical tests (e.g. Kruskal-Wallis), especially given the noise of the methods and the inter-mouse variability.

m=3 mice in miRNA microarray. Here, the internal control embedded in microarray method helps with inter-mouse variability and allowed us to use 3 mice.

Data exclusions: two populations (out of 40) were excluded from RNA-seq, due to insufficient number of reads. Number of reads threshold was set a-priori, in line with sequencing depth limitations. Stated in the Methods section.

Replication: We have used 3-5 biological replicate for each experiment - all attempts at replication were successful as there was no sample which was excluded due to outliers.

Randomization: No allocation for experimental groups in our study. All five mice underwent same procedure.

Blinding: N/A - all mice were treated the same.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a Involved in the study
- Antibodies
  - Eukaryotic cell lines
  - Palaeontology
  - Animals and other organisms
  - Human research participants
  - Clinical data

Methods

- n/a Involved in the study
  - ChIP-seq
  - Flow cytometry
  - MRI-based neuroimaging

Antibodies

**Antibodies used**

- Alexa Fluor 647 rat anti-mouse CD73 (BD, cat: 561543, lot: 7128541, clone: TY23, dilution: 1:500)
- Alexa Fluor 555 mouse anti-E-cadherin (BD, cat: 560064, lot: 6279687, clone: 36/Ecadherin, dilution: 1:500)
- Aldolase b (Thermo-Fisher, cat: P55-30218, lot: UD2748636B, polyclonal, dilution: 1:500)
- Glutamine Synthetase (Thermo-Fisher, cat: MA5-27749, lot: UD2749352F, clone: GT1055, dilution: 1:200)
- Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody conjugated to Cy5 (Thermo-Fisher, cat: A10523, dilution: 1:500)
- Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody conjugated to Cy3 (Thermo-Fisher, cat: A10521, dilution: 1:200)
- HNF4a antibody (R&D, cat: PPH1415-00, lot: A2, dilution: 1:350)
- Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, conjugated to Alexa Fluor 594 (Thermo-Fisher, cat: A21203, Ref: A21203, dilution: 1:700)
- PE-anti-E-cadherin (BioLegend, cat: 147304, lot: B238754, clone: DECMA-1, dilution: 1:300)
- APC-anti-CD73 (BioLegend, cat: 127210, lot: 8265946, clone: TY11, dilution: 1:300)
- PE-Cy7-anti-CD31 (BioLegend, cat: 204218, lot: 8212262, clone: 390, dilution: 1:300)
- APC-Cy7-anti-CD45 (BioLegend, cat: 103116, lot: 8257634, clone: 30-F11, dilution: 1:300)
Validation
Alexa Fluor 647 rat anti-mouse CD73: validated in https://doi.org/10.1016/j.cell.2018.08.063, antibody registry: AB_11218786
Alexa Fluor 555 mouse anti-Ecadherin: validated in previous refs, reported in manufacturer’s page: http://wwwbdbiosciences.com/eu/applications/research/stem-cell-research/cancer-research/human/alexa-fluor-555-mouse-anti-e-cadherin-36e-e-cadherin/p/560064
Aldolase b: validated in DOI: 10.1016/j.cmet.2018.04.003
Glutamine Synthetase: validated in the supplier’s page for the same animal (mouse) and on the same method (IF): https://www.thermofisher.com/antibody/product/Glutamine-Synthetase-Antibody-clone-GT1055-Monoclonal/MA5-27749
HNF4A: validated in DOI: 10.1016/j.jhep.2019.02.003
All FACS antibodies were validated using FMO experiments.

Animals and other organisms
Policy information about studies involving animals ARRIVE guidelines recommended for reporting animal research

Laboratory animals C57BL/6/OlaHsd Three-months old male mice, obtained from Envigo laboratories (Israel)
Wild animals the study did not involve wild animals.
Field-collected samples the study did not involve field collected samples
Ethics oversight Mouse experiments were approved by the Institutional Animal Care and Use Committee of the Weizmann Institute of Science and performed in accordance with institutional guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry
Plots
Confirm that:
☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
☒ All plots are contour plots with outliers or pseudocolor plots.
☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology
Sample preparation Liver Perfusions and hepatocytes dissociation
Five mice were anaesthetized and livers were perfused as previously previously described, with a few adjustments. A 27G syringe, connected to the perfusion line and pump, was inserted into the vena cava. 25ml of pre-warmed to 37oC EGTA buffer followed by 25ml of pre-warmed to 37oC EBS buffer with 2.3U of Liberase Blendzyme 3 recombinant collagenase (Roche Diagnostics) were cannulated into the vena cava. Shortly after the beginning of the perfusion, the portal vein was cut to allow drainage of the blood.
After perfusion, livers were explanted into a Petri dish with 25ml of pre-warmed EBS and gently minced using forceps.
Dissociated liver cells were collected and filtered through a 100um cell strainer. Cells were spun down at 30rcf for 3 min at 4oC to get hepatocytes enriched pellet. Pellet was resuspended in 25ul cold EBS.
Cells Staining
To discard dead hepatocytes, 22.5ml Percoll (Sigma) mixed with 2.5ml 10x PBS was added to the cells. Cells were centrifuged at 600rpm for 10 minutes. Supernatant containing the dead cells was aspirated and cells were resuspended in pre-warmed Hoeschst buffer (DMEM + 10% FBS + 10mM Hepes). After counting, concentration was adjusted to 2x106 cells in 1ml. To determine ploidy of hepatocytes, DNA was stained with Hoeschst (15ug/ml). Resperine (5uM) was also supplemented to the cells to prevent Hoeschst expulsion from the cells. Cells were incubated 30min at 37oC. Hepatocytes were centrifuged for 5min in 1000rpm at 4oC and supernatant was discarded. Next, cells were stained with Alexa fluor 488 Zombie green (BioLegend) to later enable the detection of viable cells by FACS. Cells were resuspended in cold PBS in a concentration of 106 cells in 100ul. Zombie-green was added in a dilution of 1:50. Cells were kept in a rotator in the dark at room temperature for 15min. After spinning down (1000rpm, 5min, 40C), cells were resuspended in FACS buffer (2mM EDTA pH 8 and 0.5% BSA in 1xPBS), in a concentration of 106 cells in 100ul. Cells were stained with PE-anti-E-cadherin (BioLegend, cat: 147304), APC-anti-CD73 (BioLegend, cat: 127210), PE-Cy7-anti-CD31 (BioLegend, cat: 102418) and APC-Cy7-anti-CD45 (BioLegend, cat: 103116), in a dilution of 1:300. FcX blocking solution (BioLegend) was added in a dilution of 1:50.

Instrument SORP-FACSAriaII sorter (BD)
Software BD FACSDiva™ Software | BD Biosciences-US
Cell population abundance Abundances of sorted populations were 2%-6% from the total recorded events. Purity of the sorted populations was validated computationally by comparing RNA-seq data of the sorted populations with positive and negative controls.
Gating strategy Cells were sorted by SORP-FACSAriaII sorter (BD) using a 130 μm nozzle and 1.5 natural density (ND) filter. Lasers compensation
Gating strategy

was corrected manually. In order to collect eight populations, each enriched with spatially-stratified hepatocytes with equal viability and ploidy levels, events were screened through the following five nested gates (Fig. 3a-b): (1) hepatocytes gate from all events – set by plotting FSC-A against SSC-A and excluding large clusters and small debris; (2) singlets FSC – set by excluding the margins of FSC-A and FSC-W plot; (3) singlets SSC – excluding upper margins of SSC-W when plotted against SSC-A; (4) live cells gates according to the Zombie-488 negative cells, comparable to unstained cells; (5) hepatocytes only, by depleting CD31 and CD45 (markers of NPCs), and (6) tetraploid hepatocytes, inferred by Hoechst histogram (Fig. 3a-b, Supplementary Fig. 1b-c). Hepatocyte size and overall protein content scale with ploidy, thus creating spurious correlations between the zonated surface markers (Supplementary Fig. 1). Sorting without ploidy stratification would result in inclusion of hepatocytes from different lobule layers, reducing spatial accuracy (Supplementary Fig. 1).

We then plotted PE-intensity for E-cadherin staining and APC-intensity for CD73 staining. The positively stained cells were determined by measuring the intensities for unstained cells. The highest intensity for unstained cells was the threshold for the positively stained cells. Each population, CD73 positive and E-cadherin positive, was further gated to four equal subpopulations, representing graded intensities of the marker. Thus, subpopulations 1, 2, 3 and 4 had equal amount of events, 1 had the highest APC-CD73 intensity while 2, 3, 4 had gradually decreasing intensities of APC. Likewise, subpopulations 5, 6, 7 and 8 were equally distributed, 8 having the highest PE-E-cadherin intensity while 7,6,5 had gradually decreasing PE intensities. Populations 4 and 5 contained cells from below positive intensity threshold, to accurately resemble mid lobule hepatocytes, in which both CD73 and E-cadherin abundances are very low (Fig. 2). All gates were set for each of the five experiments independently, with a large overlap.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.