The mammalian liver consists of hexagon-shaped lobules that are radially polarized by blood flow and morphogens1–4. Key liver genes have been shown to be differentially expressed along the lobule axis, a phenomenon termed zonation5,6, but a detailed genome-wide reconstruction of this spatial division of labour has not been achieved. Here we measure the entire transcriptome of thousands of mouse liver cells and infer their lobule coordinates on the basis of a panel of zonated landmark genes, characterized with single-molecule fluorescence in situ hybridization7. Using this approach, we obtain the zonation profiles of all liver genes with high spatial resolution. We find that around 50% of liver genes are significantly zonated and uncover abundant non-monotonic profiles that peak at the mid-lobule layers. These include a spatial order of bile acid biosynthesis enzymes that matches their position in the enzymatic cascade. Our approach can facilitate the reconstruction of similar spatial genomic blueprints for other mammalian organs.

The liver is a heterogeneous tissue that consists of hepatocytes operating in repeating anatomical units termed lobules. The hexagonally shaped lobules are composed of about 15 concentric layers of hepatocytes1–4. Blood flowing from portal veins and hepatic arteries at the corners of the lobules towards draining central veins creates gradients of oxygen, nutrients and hormones. In addition, Wnt morphogens secreted from the central vein generate an inverse polarizing field5. In line with this graded microenvironment, different radial layers sub-specialize in distinct processes, a phenomenon termed ‘liver zonation’6,8. This spatial division of labour is thought to confer optimality for liver function. For example, the outer highly oxygenated peripheral portal lobule layers express higher levels of enzymes involved in energy-demanding tasks such as gluconeogenesis and ureagenesis, whereas the inner pericentral layers specialize in glycolysis and xenobiotic metabolism.

Zonation patterns of different liver functions have been traditionally studied using RNA in situ hybridization and immunohistochemistry5,6. These techniques are limited in their sensitivity and throughput. High-throughput methods identified genome-wide differences in expression between cells enriched for portal and central hepatocytes; however, the spatial resolution of these studies was limited to two populations, isolated by either perfusion techniques6 or laser capture microdissection10.

Whole-genome reconstruction of liver zonation requires a technique to measure the entire transcriptome and the lobule coordinates of many liver cells. Single-cell RNA sequencing (scRNA-seq) enables measuring of the genome-wide expression patterns of thousands of cells11–14; however, this technique requires tissue dissociation, thus losing the spatial context of each cell. Combining scRNA-seq with in situ expression measurements of landmark genes can enable inference of the original tissue coordinates. Such an approach has recently been applied to dissect spatial gene expression signatures during embryogenesis15,16. Liver tissue is particularly challenging, as most liver genes exhibit spatially graded rather than binary expression patterns. We have previously used single-molecule fluorescence in situ hybridization (smFISH) to measure the mRNA content of hepatocytes in intact mouse liver7. This technique has the sensitivity and dynamic range to measure precise mRNA distributions at distinct lobule coordinates. We therefore combined smFISH with scRNA-seq to achieve a genome-wide reconstruction of liver zonation (Fig. 1).

As a panel of landmark genes, we chose six highly expressed liver genes with diverse zonation patterns—the pericentral genes Glul and Cyp2e1 (ref. 9) and the periportal genes Ass1 (ref. 10), Alb (ref. 9) and Cyp2f2 (ref. 9) (Fig. 2a, Extended Data Fig. 1). We segmented thousands of liver cells spanning the porto–central lobule axis of fasted mice, and reconstructed the spatially dependent probability of observing cells with distinct mRNA levels (Fig. 2b, Supplementary Information). In parallel, we used massively parallel single-cell RNA-seq (MARS-seq)17 to measure the entire transcriptome of more than 1,500 dissociated liver cells (Supplementary Table 1). Our measurements revealed three distinct populations consisting of Kupffer cells, endothelial cells and hepatocytes (Extended Data Fig. 2a–c). Hepatocytes were highly heterogeneous in their expression patterns and exhibited clear gradients that matched our landmark gene profiles (Fig. 3a, b). The urea cycle gene Ass1 and the glutamine synthetase gene Glul were expressed in a mutually exclusive manner in the single-cell data, demonstrating the purity of the single-cell isolation (Extended Data Fig. 2d, e).

To combine the scRNA-seq results with our smFISH-obtained zonation profiles we divided the porto–central lobule axis into nine layers and developed a probabilistic inference algorithm to compute the likelihood of each cell belonging to any of these layers, based on the expression of our panel of landmark genes (Extended Data Fig. 3, Supplementary Information, Supplementary Tables 2, 7). Our reconstruction accuracy approached saturation at six genes. It was strongly dependent on the extent of zonation of our landmark genes, and only weakly dependent on the intra-layer cell-to-cell variability (Extended Data Fig. 4). We validated the precision of our reconstructed zonation profiles using smFISH on 20 genes with diverse profiles and found an excellent overall correspondence between the predicted and measured profiles (Extended Data Fig. 5).

Our analysis revealed that around 50% of the expressed liver genes are non-randomly spatially zonated (Fig. 3c, Supplementary Table 3; 3,496 of 7,277 genes with q value <0.2, Kruskal–Wallis test), an order of magnitude higher than previous estimates3. Our reconstructed profiles recapitulated the expression of genes over more than four orders of magnitude, including genes with low expression, such as the pericentral Axin2 (ref. 2) and the periportal Sox9 (ref. 18) (Fig. 3c, Extended Data Fig. 6). The porto–central ratio of our reconstructed profiles correlated with previous datasets that compared the transcriptome of pericentral and periportal hepatocytes2,9,10 (Extended Data Fig. 7).

1Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. 2Department of Immunology, Weizmann Institute of Science, Rehovot, Israel. 3Biological Services, Weizmann Institute of Science, Rehovot, Israel. *These authors contributed equally to this work.
We next examined which signalling pathways shape the global liver zonation profiles that our method revealed. Diffusible Wnt signals, originating at the pericentral endothelial cells, have been suggested to be of major importance in inducing pericentral zonation profiles. We found that liver genes that increased in expression in Wnt-hyperactivating liver-specific Apc knockout (Apc-KO) mice were predominantly pericentral (810 of our 3,496 zonated genes, median peak expression at layer 1), whereas genes that decreased in Apc-KO mice had a significantly stronger periporal bias (193 of our 3,496 zonated genes, median peak expression at layer 6, Wilcoxon rank-sum, \( P < 10^{-31} \), Fig. 3d, Supplementary Table 4). Similarly, 95 of the 3,496 zonated genes that were shown to increase in expression upon chronic hypoxia were significantly biased towards the low-oxygenated pericentral layers, compared to 45 of the zonated genes that decreased in chronic hypoxia (Fig. 3e, Extended Data Fig. 8a, b, \( P = 0.022 \)). Thus, Wnt signalling and low oxygen are major factors inducing pericentral zonation profiles.

Ras signalling has been hypothesized to induce peripoportal zonation profiles. We found that the zonation profiles of genes that increased in expression in Ha-ras hyperactivated tumours were significantly more peripoportal compared to genes that reduced in expression (\( P = 0.0001 \), Fig. 3f, Extended Data Fig. 8c, d). We also found that pituitary hormones repress periporal genes, as evident by the periporal profile of genes that increase in hypopituitary dwarf mice (\( P = 0.0054 \), Fig. 3g, Extended Data Fig. 8e, f). Importantly, about two thirds of the zonated genes (2,314 out of 3,496 genes) were not predicted targets of either Wnt, hypoxia, Ras signalling or pituitary hormones, suggesting that their zonation is affected by combinatorial regulation of these factors or by additional, yet to be identified, morphogens or blood-borne factors.

To explore the zonation patterns of specific biological pathways, we analysed the KEGG database and found that 25 of the 186 KEGG pathways were enriched for zonated genes (hypergeometric test, \( q < 0.1; \)).
Supplementary Table 5, Extended Data Fig. 9a). The oxidative phosphorylation pathway exhibited higher expression in the periportal layers, where oxygen concentration is higher, as were the secreted proteins composing the complement and coagulation cascades pathway. More generally, we identified a periportal bias in the mRNA of genes that encode liver-secreted proteins (Extended Data Fig. 9b). Allocating the ATP-demanding task of plasma protein production to the well-oxygenated periportal layers may be energetically efficient, as an estimated 20% of liver oxygen consumption is dedicated to this task. Pericentrally biased pathways included detoxification pathways such as xenobiotic metabolism and glutathione metabolism (Extended Data Fig. 9c), as well as bile acid biosynthesis and proteasome components (Extended Data Fig. 9a).

All liver zonation profiles previously described were monotonically increasing or decreasing porto-centrally. Our high spatial resolution enabled identification of a new class of non-monotonic zonation profiles that peak at intermediate lobule layers (Figs. 3c, 4a, Extended Data Figs 5, 10). Although there was no significant gene ontology annotation associated with these genes, they included key liver genes such as Hamp and Hamp2 that encode hepcidin, a secreted liver hormone that regulates systemic iron levels (Fig. 4a). Additional non-monotonic genes included Igfbp2, Mup3 and Cyp8b1 (Fig. 3c, Extended Data Figs 5, 10). The non-monotonic expression of these genes could not be explained by the previously identified non-monotonic pattern of liver polyploidy (Extended Data Fig. 10e, Methods).

Both Igfbp2 and Cyp8b1, peaking at the mid-lobule layers, participate in pathways that consist of sequential genes, which we found to be expressed in sequential lobule coordinates. Igf1 is an abundant circulating growth factor with multiple roles in regulating organismal physiology. The secreted binding proteins Igfbp stabilize Igf1.
and limit its binding to Igf receptors26. We found Igf1 to be highly expressed in the periportal layers and its binding proteins Igfbp2 and Igfbp1 expressed in the mid-lobule and pericentral layers, respectively (Extended Data Figs 5, 10c, d). This spatial order may expose a feedback system, wherein Igfbp production matches the levels of upstream-secreted Igf, or could be related to other non-endocrine functions of Igfbp.

Liver bile acids are produced in pericentral hepatocytes from cholesterol, consequently flowing through bile canaliculi towards the draining bile duct5. The neutral pathway of bile acid biosynthesis27, consisting of the core sequence Cyp7a1→Hsd3b7→Cyp8b1→Cyp27a1, exhibited a spatial order that matched the position in the enzymatic cascade (Fig. 4b, c). Whereas Cyp7a1 and Hsd3b7 were most abundant in the pericentral layer 1, the next enzyme, Cyp8b1, peaked in layers 2–3 (Fig. 4b, c, Extended Data Fig. 10f). Baat, the gene encoding bile acid-CoA:amino acid N-acyltransferase, an enzyme that conjugates bile acids before their secretion, was also depleted in layer 1 (Fig. 4b). The lower levels of Cyp8b1 in layer 1 compared to layer 2 may potentially indicate that intermediates are transferred between layers 1 and 2. Alternatively it could suggest that Cyp7a1 and Hsd3b7 harbour additional functions in layer 1 or that Cyp8b1 harbours additional functions in layer 2. Using the database of KEGG metabolic networks we identified additional zonated liver enzyme pairs that are significantly spatially discordant and either operate on the same substrate, produce the same product, or act sequentially such that the product of one enzyme is the substrate of the other (Supplementary Table 6). This constitutes a resource for further exploration of liver division of labour.

Our study challenges the traditional binary classification of liver into periportal and pericentral hepatocytes and reveals multiple roles for the intermediate lobule coordinates. The broad spatial heterogeneity revealed here may represent an optimal balance between the diverse liver tasks28. Our work provides the resource to explore the design principles governing this global spatial division of labour. The detailed spatial information of liver gene expression presented here can serve as an important tier in efforts to achieve comprehensive tissue-level metabolic reconstructions29.

Liver gene expression is highly dynamic, with multiple genes showing circadian and metabolic variations30. Although our study focused on zonation profiles of fasted mice, similar reconstructions will decipher the dynamics of liver zonation profiles along these temporal axes. Another natural extension is the analysis of liver pathologies. Reconstruction of liver zonation profiles in response to liver intoxication will expose the plasticity of zonation and the extent to which different layers compensate for the functional loss of other layers. Our
approach for combining smFISH and scRNA-seq can be readily applied to extract spatial blueprints of other structured mammalian organs in diverse physiological and pathological states.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 June; accepted 19 December 2016. Published online 6 February; corrected online 15 February 2017 (see full-text HTML version for details).

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. Benhamouche, S. et al. Apc tumor suppressor gene is the “zonation-keeper” of mouse liver. Dev. Cell 10, 793–797 (2006).
4. Coint, S. & Perret, C. in Molecular Pathology of Liver Diseases (ed. Monga, S. P. S.) 7–16 (Springer US, 2011).
5. Gebhardt, R. Metabolic zonation of the liver: regulation and implications for liver function. PharmacoI. Ther. 93, 275–394 (1992).
6. Jungermann, K. & Kietzmann, T. Zonation of parenchymal and nonparenchymal metabolism in liver. Annu. Rev. Nutr. 16, 179–203 (1996).
7. Bahar Halpern, K. et al. Bystay gene expression in the intact mammalian liver. Mol. Cell 58, 147–156 (2015).
8. Gebhardt, R. & Matz-Soja, M. Liver zonation: novel aspects of its regulation and its impact on homeostasis. World J. Gastroenterol. 20, 8491–8504 (2014).
9. Braeuning, A. et al. Differential gene expression in periportal and perivenous mouse hepatocytes. FBSJ 273, 5051–5061 (2006).
10. Saito, K., Negishi, M. & James Squires, E. Sexual dimorphisms in zonal gene expression in mouse liver. Biochem. Biophys. Res. Commun. 436, 730–735 (2013).
11. Hashimshony, T., Wagner, F., Sher, N. & Yani, I. CEL-seq: single-cell RNA-seq by multiplexed linear amplification. Cell Reports 2, 666–673 (2012).
12. Wu, A. R. et al. Quantitative assessment of single-cell RNA-sequencing methods. Nat. Methods 11, 41–46 (2014).
13. Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347, 1138–1142 (2015).
14. Grün, D. et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525, 251–255 (2015).
15. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502 (2015).
16. Achim, K. et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. Nat. Biotechnol. 33, 503–509 (2015).
17. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343, 776–779 (2014).
18. Font-Burgada, J. et al. Hybrid periportal hepatocytes regenerate the injured liver without giving rise to cancer. Cell 162, 766–779 (2015).
19. Gougelet, A. et al. T-cell factor 4 and β-catenin chromatin occupancies pattern zonal liver metabolism in mice. Hepatology 59, 2344–2357 (2014).
20. Baze, M. M., Schlauch, K. & Hayes, J. P. Gene expression of the liver in response to chronic hypoxia. Physiol. Genomics 41, 275–296 (2010).
21. Hailfinger, S., Jaworski, M., Braeuning, A., Buchmann, A. & Schwarz, M. Zonal gene expression in murine liver: lessons from tumors. Hepatology 43, 407–414 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Milo and all members of our laboratory for valuable comments. We thank M. Schwarz, A. Braeuning and S. Coint for sharing their data and A. Sharp, E. Ariel, H. Gagzi and J. Gavish for help with experimental procedures. I.A. is supported by the European Research Council (309788), and the Israel Science Foundation, the Ernst and Bonnie Beutler Research Program of Excellence in Genomic Medicine and the Helen and Martin Kimmel award for innovative investigation. I.A. is the incumbent of the Alan and Laraine Fischer Career Development Chair. S.I. is supported by the Henry Chanoch Krenter Institute for Biomedical Imaging and Genomics, The Leir Charitable Foundations, Richard Jakubskind Laboratory of Systems Biology, Cymerman-Jakubskind Prize, The Lord Sief of Brimpton Memorial Fund, the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (grants 1902/12 and 1796/12), the Israel Science Foundation grant No. 1486/16, the EMBO Young Investigator Program and the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement number 335120. S.I. is the incumbent of the Philip Harris and Gerald Ronson Career Development Chair.

Author Contributions K.B.H., B.T., M.G., S.L., E.E.M., S.B. and A.S.A. performed smFISH experiments. K.B.H., B.T. and O.M.-N. performed the single-cell isolation and FACS sorting. O.M.-N. performed the experimental scRNA-seq procedures, A.G. and E.D. implemented the scRNA-seq computational protocols, R.S. and S.I. developed and implemented the spatial inference algorithms, R.S., D.L. and S.I. performed data analysis, A.E.M. and A.B. performed the GypB1 validation experiments, I.A. and S.I. supervised the study. I.A. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this paper. Correspondence and requests for materials should be addressed to I.A. (ido.amit@weizmann.ac.il) or S.I. (shalev.itzkovitz@weizmann.ac.il).

Reviewer Information Nature thanks K. Kaestner, A. Raj and the other anonymous reviewer(s) for their contribution to the peer review of this work.
**METHODS**

**Mice and tissues.** All animal studies were approved by the Institutional Animal Care and Use Committee of WIS. C57BL/6 male mice aged 2 months were housed under reverse-phase cycle, and fasted for 5h starting at 07:00. All mice were anaes-
thetized with an intraperitoneal injection of a ketamine (100 mg·kg⁻¹) and xylazine (10 mg·kg⁻¹) mixture. For smFISH, liver tissues were collected and fixed in 4% paraformaldehyde for 3h; incubated overnight with 30% sucrose in 4% paraformal-
dehyde and then embedded in OCT. 7μm cryosections were used for hybridization. Mouse liver cells for RNA-seq were extracted from four mice and each smFISH result was based on at least 2 mice. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. No statistical methods were used to predetermine sample size.

**Hybridization and imaging.** Probe library constructions, hybridization proce-
dures and imaging conditions were previously described12,32. All smFISH probe libraries (Supplementary Table 8), with the exception of glutamine synthetase (Glul), were coupled to Cy5. To detect cell borders, alexa fluor 488 conjugated phallolidin (Rhenum A12379) was added to the GLOX buffer wash12. Portal node was identified morphologically on DAPI images on the basis of bile ductule, central vein was identified using smFISH for Glud in TMR, included in all hybridizations. For zonation profiles, images were taken as scans spanning the portal node to the central vein. To compute single-cell mRNA concentration of our landmark genes, we used imageM32 to detect dots, extracted the sum of all dot intensities of the smFISH signal for each segmented cell within the first 3μm of the Z-stack, and divided by the segmented cell volume to obtain the intensity concentration per cell. As the landmark genes were highly abundant, dots often coalesced, leading to underestimates of the true cellular gene expression when counting dots, whereas the sum of dot intensities was found to better capture the full dynamic range33. The gene expression distributions of the landmark genes were based on at least 800 cells from at least 10 lobules and 2 mice per gene. To validate the predicted zonation, we imaged 20 additional genes on at least 10 lobules and 2 mice per gene (Extended Data Fig. Sb). Profiles for Lgfbp1, Cyp27al1, Glucl, Cypb51b1, Igf2rb, Pck1, Cpsi, Arg1, G6pc, Uox, Igf1, Pigr and Acly were generated by counting dots and dividing the number of dots in radial layers spanning the portal–central axis by the layer volume. Profiles for Cyp1a2, Rnasee4, Gsta3, Ugt1a1, Hmp, Map3 and Apoa1 were generated by quantifying the average background-subtracted intensity of the smFISH images at sequential lobule layers.

**Hepatocytes isolation.** Mouse hepatocytes were isolated by a modification of the two-step collagenase perfusion method of Seglen34 from 5 h fasted, 2-month-old male C57BL/6 mice. Single cells were isolated from four mice. Digestion step was performed with Liberase Blendzyme 3 recombinant collagenase (Roche Diagnostics) according to the manufacturer's instruction. Isolated hepatocytes were taken directly to sorting.

**Single-cell sorting.** Cells were sorted with SORP-FACS AriaII machine using a 130μm nozzle. Dead cells were excluded on the basis of 5 μg·ml⁻¹ propidium iodide (Invitrogen) incorporation. Cells adhering to each other (that is, doublets) were eliminated on the basis of pulse width. We used a 25,000–250,000 FSC-A gate. For three of the mice a 1.5 neutral density (ND) filter and was used to enrich non-parenchymal cells. Hepatocytes were sorted into 384-well cell capture plates using an automated pipeline. The liver secretome (Extended Data Fig. 9b) was based on ref. 37. To identify spatial division of labour within specific pathways we examined the 62 mouse metabolic pathway maps from KEGG database35. We extracted from each pathway map pairs of enzymes that have either a shared substrate, a shared product, or pairs in which a product of one enzyme is the substrate of the second enzyme. Upon obtaining all pairs we further selected those in which both enzymes are expressed in the liver (mean expression higher than 5 x 10⁻⁶ UMI per cell) and are signifi-
cantly zonated (Kruskal–Wallis q < 0.2). For each such pair we assigned a score that reflects the spatial discordance in the zonation profiles of the both enzymes: $S(E_i, E_j) = (E_i(x_i) - E_j(x_j)) \times (E_i(x_i) - E_j(x_j))$ Where $E_i(x)$ is the mean-normalized zonation profile of enzyme $i$ and $x_i$ is the layer at which $E_i(x)$ peaks. Negative scores indicate that the two enzymes peak at different layers, and the quantity of the score reflects the expression differences between the two layers. Spatial interaction between two enzymes that peak at the same layer has a score of zero. In addition, for every triplet of connected enzymes $E_1 \rightarrow E_2 \rightarrow E_3$ we also included the indirect pair $E_1 - E_3$. To produce a concise set of significantly zonated spatially discordant enzyme pairs we randomly drew 100,000 pairs of enzymes among all expressed liver enzymes (731 genes) and recomputed their scores. For our concise set we selected pairs with a score lower than the 10th percentile score among the randomized set of pairs (0.14, Supplementary Table 6). Although the direction of flow of metabolites is hard to systematically determine for these pathways (pericentral direction through blood versus periportal direction through bile), this database serves as a resource for future focused exploration of liver spatial division of labour of individual pathways.

**Algorithm for spatial reconstruction of zonation profiles.** The algorithm for spatial reconstruction of zonation profiles is described in detail in the Supplementary Information. Matlab code used for the inference is available upon request.
Data availability. Data generated during this study have been deposited in Gene Expression Omnibus (GEO) with the accession code GSE84498. Data referenced in Fig. 3d and Extended Data Fig. 7a are available on request from the authors. Data referenced in Fig. 3e and Extended Data Fig. 8a, b are available in supplementary table 1 of ref. 20. Data referenced in Fig. 3f and Extended Data Fig. 8c, d are available in supplementary table 4b of ref. 22. Data referenced in Fig. 3g and Extended Data Fig. 8e, f are available in GEO with the accession code GSE3129 (ref. 23); data referenced in Extended Data Fig. 7b are available in GEO with the accession code GSE49707 (ref. 10); data referenced in Extended Data Fig. 7c, d are available in GEO with the accession code GSE68806 (ref. 2). Data referenced in Extended Data Fig. 9b are available in table 6 of ref. 37.

31. Itzkovitz, S. et al. Single-molecule transcript counting of stem-cell markers in the mouse intestine. Nat. Cell Biol. 14, 106–114 (2011).
32. Lyubimova, A. et al. Single-molecule mRNA detection and counting in mammalian tissue. Nat. Protocols 8, 1743–1758 (2013).
33. Bahar Halpern, K. et al. Nuclear retention of mRNA in mammalian tissues. Cell Reports 13, 2653–2662 (2015).
34. Seglen, P. O. Preparation of rat liver cells. 3. Enzymatic requirements for tissue dispersion. Exp. Cell Res. 82, 391–398 (1973).
35. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360 (2015).
36. Eden, E., Navon, R., Steinfeld, L., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48 (2009).
37. Zhang, Y. et al. Strategy for studying the liver secretome on the organ level. J. Proteome Res. 9, 1894–1901 (2010).
38. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30 (2000).
Extended Data Figure 1 | Low magnification images of the six landmark genes. smFISH examples of our 6 landmark genes—the pericentral genes *Glul* and *Cyp2e1* and the periportal genes *Ass1*, *Asl*, *Alb* and *Cyp2f2*. Bright cells have high mRNA content. Scale bars, 100 μm. CV, central vein; PN, portal node. Micrographs are representative of at least ten lobules and at least two mice per gene.
Extended Data Figure 2 | Single-cell RNA-seq reveals three distinct liver cell populations. a–c, t-SNE plots. Each dot is a cell coloured according to the aggregated expression of the hepatocyte marker genes Aapo1, Glul, Acly, Asl, Cyp2e1, Cyp2f2, Ass1, Alb, Mup3, Pck1, G6pc (a), the Kupffer cell marker genes Irf7, Spic, Clec4f (b) and the endothelial cell markers Ushbp1, Myf6, Oii3, Il1a, F8, Bmp2, C1qtnf1, Mmnrn2, Pcdh12, Dpp4 (c). d, smFISH of a liver lobule demonstrating the antagonistic expression of Glul (red dots), and Ass1 (green dots). Scale bar, 20 μm. Inset of highlighted area demonstrates the minimal co-expression of these two genes. Scale bars, 5 μm. Blue, DAPI-stained nuclei. CV, central vein; PN, portal node. Micrographs are representative of at least 10 lobules and at least two mice per gene. e, Ass1 and Glul are mutually exclusive in the single-cell RNA-seq data. Each dot is a cell, x axis is the expression of Glul in fraction of total cellular UMI, y axis is the expression of Ass1.
Extended Data Figure 3 | Prior and posterior probability computation of the scRNA-seq data. a, b, Zone-dependent probabilities of observing different expression levels for each of the six landmark genes. Horizontal lines denote the UMI observed for each gene for cell 609 (a) and cell 629 (b). The posterior probabilities for each layer suggest that cell 609 most probably originated from the periportal layer 8, whereas cell 629 originated from the pericentral layer 1. Note that the posterior probabilities incorporate the lower sampling of cell 609, which had a total of 3,358 UMI, compared to cell 629, which had 9,913 UMI. c, d, Prior probability of observing a hepatocyte from each of the 9 layers. c, Hexagonal lobule geometry and 9 concentric circles at constant radii increments spanning the central vein and portal node. d, Area of each layer is the intersection with the hexagon.
Extended Data Figure 4 | Sensitivity of reconstruction to number and features of landmark genes. a, Zonation properties of our landmark genes. $-\log_{10}$ of the Wilcoxon rank-sum P values for the comparison of smFISH expression measurements in cells in sequential lobule layers for each of the 6 landmark genes. Stars denote layer pairs with significant changes in expression ($P$ value < 0.05). b, Mean reconstruction accuracy for different subsets of landmark genes (defined as 1 minus the mean of the sum-squared differences between the profiles predicted using each subset and the profiles obtained when using the entire panel of landmark genes). Dot colour represents the number of landmark genes in each subset. c, Reconstruction accuracy approaches saturation when using 6 landmark genes. G, Glul; Cf, Cyp2f2; Ce, Cyp2e1; As, Ass1; Al, Asl; Ab, Alb. Dashed line connects the most accurate partial subset of landmark genes for each panel size. d, The contribution of each landmark gene to zonation reconstruction is strongly dependent on its spatial non-uniformity among different layers and less on its intra-layer variability. x axis shows the entropy of the summed-normalized average profile of each landmark gene. y axis shows the average among all layers of the coefficient of variation of the intra-layer expression levels. The score for each gene is the average of the ratio in reconstruction error with and without the gene among all combinations that include the gene. High scores indicate that the landmark gene strongly improves reconstruction when added to combinations of other landmark genes.
Extended Data Figure 5 | Validation of reconstructed zonation profiles using smFISH. 

a, Reconstructed zonation profiles based on the scRNA-seq data (blue) and smFISH measurements (red) for our landmark genes. 
b, Validation of reconstructed profiles for 20 genes not used for the inference algorithm. Profiles in a, b, are normalized by the mean, patches show s.e.m. Lobule layers are numbered from the central vein (layer 1) to the portal node (layer 9). Note that Pck1 has a broader profile compared to other studies, since our measurements were done on fasted mice, a metabolic state in which gluconeogenesis becomes substantial in pericentral layers.
Extended Data Figure 6 | Reconstructed zonation profiles capture a wide dynamic range. Reconstructed zonation profiles of the pericentral gene Oat (purple) and the pericentral progenitor marker Axin2 (red), and the periportal urea cycle enzyme gene Arg1 (green) and the periportal progenitor marker Sox9 (blue). Dashed box highlights a blow-up of Axin2 and Sox9, genes with 100-fold lower expression than Oat and Arg1. Expression values are the estimated fraction of the total cellular mRNA molecules. Lobule layers are numbered from the central vein (CV, layer 1) to the portal node (PN, layer 9). The slightly broader zonation profile of Sox9 mRNA compared to the Sox9 signal observed in18 using Sox9–GFP or Sox9–CreERT-R26RtdTomato may stem from differences in the dynamic range of the detection methods or from differences in the mouse metabolic states (fasted versus ad libitum).
Extended Data Figure 7 | Portocentral ratio of reconstructed zonation profiles correlates with previous studies. a, Correlation between the pericentral bias computed from our data and the one from ref. 9. y axis is our pericentral bias (PC/PP), computed as the ratio between the median zonation profiles in layers 1–4 and the median in layers 6–9. x axis is the ratio between expression in pericentrally and periportally enriched genes from ref. 9. 79 of the 88 genes considered pericentral according to Braeuning et al. (red circles) are also pericentrally biased in our dataset (hypergeometric test \( P < 4 \times 10^{-9} \)). 70 of the 81 genes considered periportal according to Braeuning et al. (green circles) are also periportally biased in our dataset (hypergeometric test \( P < 1 \times 10^{-16} \)). Black squares mark genes that we found to be significantly zonated. b, Pericentral bias computed by our method correlates with previous bulk RNA-seq studies that compared pericentral and periportal populations isolated via laser capture microdissection\(^9\). Spearman \( R = 0.74; P < 1 \times 10^{-80} \). Black dots represent genes considered significantly zonated in ref. 10. a, b, Include genes with average expression across cells which is higher than \( 5 \times 10^{-6} \) of the total UMI counts. c, Reconstructed zonation profiles for the genes found in ref. 2 to have higher expression in Axin2\(^{+}\) pericentral hepatocytes (\( q \) value < 0.2). d, Reconstructed zonation profiles for the genes found in ref. 2 to have lower expression in pericentral Axin2\(^{+}\) hepatocytes (\( q \) value < 0.2).
Extended Data Figure 8 | Zonation profiles of genes that significantly increased or decreased when perturbing signalling pathways. 

a, b, Zonation profiles of liver genes previously shown to increase (a) or decrease (b) in liver of mice exposed to chronic hypoxia (pO2 = 11.5 kPa, compared to pO2 = 18.0 kPa). c, d, Zonation profiles of genes shown to significantly increase (c, log-fold > 2) or decrease (d, log-fold < −2) in expression in liver tumours harbouring an activating Ha-ras mutation23 (FDR-corrected P value < 0.05). e, f, Zonation profiles of genes that have higher (e, fold > 1) or lower (f, fold < 1) expression in hypopituitary mice23 (FDR-corrected P value < 0.05).
Extended Data Figure 9 | KEGG pathways enriched for zonated genes. a, Average max-normalized zonation profiles of KEGG pathways enriched for zonated genes. Panel displays all pathways with more than 10 genes and hypergeometric test q value < 0.1 (Supplementary Table 5 exhibits full results). Each profile was normalized by its maximum along all layers. b, Periportal bias of liver secreted proteins. Genes are based on an annotated liver secretome. c, Pericentral bias for liver detoxification genes. Shown are genes for cytochrome P450, Uridine 5′-diphospho-glucuronosyltransferase and glutathione S-transferase. Images in b, c include significantly zonated genes from each pathway (Kruskal–Wallis q value < 0.01 and more than 5 × 10^-5 of the total cellular UMIs on average in at least one of the 9 layers), each profile normalized to maximum of 1.
Extended Data Figure 10 | Non-monotonic zonation profiles of liver genes. a, Max-normalized zonation profiles of the concise set of 46 non-monotonic genes (Methods). b, Mup3 is highly variable but exhibits a peak at layer 7 and decreased expression in both the pericentral layer 1 and perportal layer 9. c, Igfbp2 exhibits a non-monotonic zonation profile, peaking at layers 3–6. CV, central vein; PN, portal node. d, Igfbp2, Igfbp1 and Igfbp4 are downstream to Igf1, encoding the secreted protein that they bind and stabilize. e, Non-monotonic zonation profiles observed do not stem from ploidy-specific regulation. Each dot represents a pair of adjacent cells (within 30μm of each other), x axis is the higher ploidy cell, y axis is the lower ploidy cell. P values are Wilcoxon sign-ranked tests. Gene expression for Igfbp2, Hamp and Cyp8b1 was quantified as the number of smFISH dots divided by the segmented cell volumes; Mup3 expression was quantified as the average intensity, due to its higher abundance. Quantification for each gene based on at least 60 pairs. f, Immunohistochemistry of Cyp8b1 demonstrates higher protein concentration in layers 2–3 (dashed curve) compared to layer 1 (dotted curve). Scale bar, 50μm. Micrographs are representative of seven lobules in two mice.
ERRATUM

doi:10.1038/nature21729

Erratum: Single-cell spatial reconstruction reveals global division of labour in the mammalian liver

Keren Bahar Halpern, Rom Shenhav, Orit Matcovitch-Natan, Beáta Tóth, Doron Lemze, Matan Golan, Efi E. Massasa, Shaked Baydatch, Shanie Landen, Andreas E. Moor, Alexander Brandis, Amir Giladi, Avigail Stokar-Avihail, Eyal David, Ido Amit & Shalev Itzkovitz

Nature 542, 352–356 (2017); doi:10.1038/nature21065

In this Letter, owing to an error during the production process, the Gene Expression Omnibus (GEO) accession number in the 'Data availability' section was incorrectly listed as 'GSE84490' instead of 'GSE84498'; this has been corrected online.