Space-time logic of liver gene expression at sublobular scale

Colas Droin[1+], Jakob El Kholtei[2+], Keren Bahar Halpern[2+], Clémence Hurni[1], Milena Rozenberg[2], Sapir Muvkadi[2], Shalev Itzkovitz[2*], Felix Naef[1*]

[1]Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

[2]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel

[+]These authors contributed equally to this work.

*Corresponding authors**:**

Shalev Itzkovitz, Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel

Email: shalev.itzkovitz@weizmann.ac.il; +972 8 9343104:

Felix Naef, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Email: felix.naef@epfl.ch; phone: +41 21 6931621

Abstract

The mammalian liver performs key physiological functions for maintaining energy and metabolic homeostasis. Liver tissue is both spatially structured and temporally orchestrated. Hepatocytes operate in repeating anatomical units termed lobules and different lobule zones perform distinct functions. The liver is also subject to extensive temporal regulation, orchestrated by the interplay of the circadian clock, systemic signals and feeding rhythms. Liver zonation was previously analyzed as a static phenomenon and liver chronobiology at the tissue level. Here, we use single-cell RNA-seq to investigate the interplay between gene regulation in space and time. Categorizing mRNA expression profiles using mixed-effect models and smFISH validations, we find that many genes in the liver are both zonated and rhythmic, most of them showing multiplicative space-time

effects. Such dually regulated genes cover key hepatic functions such as lipid, carbohydrate and amino acid metabolism. In particular, our data suggest that rhythmic and localized expression of Wnt targets may be explained by rhythmic Wnt signaling from endothelial cells near the central vein. Core circadian clock genes are expressed in a non-zonated manner, indicating that the liver clock is robust to zonation. Together, our comprehensive data reveal how liver function is compartmentalized spatio-temporally at the sub-lobular scale.

Introduction

The liver is a vital organ maintaining body physiology and energy homeostasis. The liver carries out a broad range of functions related to carbohydrate and lipid metabolism, detoxification, bile acid biosynthesis and transport, cholesterol processing, xenobiotics biotransformation, and carrier proteins secretion. Notably, the liver performs catabolic and anabolic processing of lipids and amino acids and produces the majority of plasma proteins[1]. Liver tissue is highly structured on the cellular scale, being heterogeneous in both cell-type composition and microenvironment[2]. In fact, liver tissue is made up of millions of repeating anatomical and functional subunits, called lobules, which in mice contain hepatocytes arranged in 12-15 concentric layers with a diameter of about 0.5mm[3]. On the portal side of the lobule, blood from the portal vein and the hepatic arteriole enters small capillaries called sinusoids and flows to the central vein. This is accompanied with gradients in oxygen concentration, nutrients and signaling along the porto-central axis, with the latter notably involving the Wnt pathway[4,5]. Due to this polarization, hepatocytes in different layers perform separate functions, a phenomenon termed *liver zonation*[1,6].

Recently, we combined single-cell RNA-sequencing (scRNA-seq) of dissociated hepatocytes and single-molecule RNA fluorescence in situ hybridization (smFISH) to reconstruct spatial mRNA expression profiles along the porto-central axis[7]. This analysis revealed an unexpected breadth of spatial heterogeneity, with ~50% of genes showing spatially non-uniform patterns. Among them, functions related to ammonia clearance, carbohydrate catabolic and anabolic processes, xenobiotics detoxification, bile acid and cholesterol synthesis, fatty acid metabolism, targets of the Wnt and Ras pathways, and hypoxia-induced genes were strongly zonated.

In addition to its spatial heterogeneity, the liver is also highly dynamic temporally. Chronobiology studies showed that temporally gated physiological and metabolic programs in the liver result from the complex interplay between the endogenous circadian liver oscillator, rhythmic systemic signals, and feeding/fasting cycles[8,9,10]. An intact circadian clock has repeatedly been demonstrated as key for healthy metabolism, also in humans[11]. Temporal compartmentalization can prevent two opposite and incompatible processes from simultaneously occurring, for example, glucose is stored as glycogen following a meal and is later released into the blood circulation during fasting period to maintain homeostasis in plasma glucose levels. Functional genomics studies of the circadian liver were typically performed on bulk liver tissue[12]. In particular, we and others showed how both the circadian clock and the feeding fasting cycles pervasively drive rhythms of gene expression in bulk, impacting key sectors of liver physiology such as lipid and steroid metabolism[13–15].

Here, we asked how these spatial and temporal regulatory programs interact on the level of individual genes and liver functions more generally. In particular, can zonated gene expression patterns be temporally modulated on a 24 h time scale? And conversely, can rhythmic gene expression patterns observed in bulk samples exhibit sub-lobular structure? More complex situations may also be envisaged, such as time-dependent zonation patterns of mRNA expression (or, equivalently, zone-dependent rhythmic patterns), or sublobular oscillations that would escape detection on the bulk level due to cancelations. On the physiological level, it is of interest to establish how hepatic functions might be compartmentalized both in space and time. To study both the spatial and temporal axes, we performed scRNA-seq of hepatocytes at four different times along the 24 h day, extending our previous approach[7,16] to reconstruct spatial profiles at each time point. The resulting space-time patterns were statistically classified using a mixed-effect model describing both spatial and temporal variations in mRNA levels. In total, ~5000 liver genes were classified based on their spatio-temporal expression profiles, and a few representative profiles were further analyzed with smFISH. Overall, this approach helped to elucidate the richness of space-time gene expression dynamics of the liver and provides a comprehensive view on how spatio-temporal compartmentalization is utilized in the mammalian liver.

Results

Single-cell RNA-seq captures spatiotemporal gene expression patterns in mouse liver

To investigate spatio-temporal gene expression patterns in mouse liver, we sequenced mRNA from liver cells obtained from 10 mice at 4 different times of the day (ZT = 0h, 6h, 12h and 18h, two to three replicates per time point). We here focused on hepatocytes by enrichment of cells according to size and *in silico* filtering, yielding a total of 19663 cells (several filtering steps are involved, Methods). To validate that the expected axes of variation are present in the scRNA-seq data, we generated a low-dimensional representation of all cells (t-SNE projections) and colored cells either by their position along the centro-portal axis (layers) (Figure 1A) or time (Figure 1B). This revealed that known portally and centrally expressed transcripts, such as *Cyp2f2* (Figure 1C) or *Cyp2e1* (Figure 1D), respectively, mark cells in opposite regions of the projections. Likewise, time-of-day expression varied along an orthogonal direction, as shown for the *Elovl3* gene peaking at ZT0 (Figure 1E).

To reduce the complexity of the spatial variation in mRNA levels, we here introduced eight different lobule layers to describe gene expression along the centro-portal axis. For this, we adapted our previous method[16], with the modification that only transcripts that did not vary across time points were used as landmark zonated genes (Methods). The resulting reconstructed mRNA expression profiles yielded 80 (8 layers over 10 mice) data points for each transcript. These reconstructions faithfully captured reference zonated genes, with both central, and portal, expression (Figure 1F), such as *Glul*, and *Ass1*, respectively. The reconstructions also included reference core clock and rhythmic output genes, such as *Bmal1* (also named *Arntl*) and *Dbp*, with large differences in expression across the time points, and peaking at the expected times (Figure 1G). Finally, genes showing both zonated and rhythmic mRNA accumulation were found (Figure 1H), with both central (*Elovl3*) or portal (*Pck1*) expression patterns, and with specific peak times. Since most of the zonated profiles showed exponential shapes, and gene expression changes typically occur on a log scale[17], we log-transformed the data for further analysis (Methods, Figure 1F and Supplementary Figure 1A).

Space-time mRNA expression profiles can be categorized according to zonation and rhythmicity

Next, we investigated if zonated gene expression patterns can be dynamic along the day, or conversely whether temporal expression patterns might be zone-dependent. To select a reliable set of reconstructed mRNA expression profiles for subsequent analyses, we first discarded lowly expressed genes, as well as genes with significant biological variability across replicate liver samples. This yielded 5058 spatio-temporal gene expression profiles (Supplementary Figure 2A). An exploratory analysis of variance clearly identified zonated genes, rhythmic genes, and fewer genes showing variability along both axes, with known zonated and rhythmic genes distributed as expected (Figure 2A).

To identify possible dependencies between spatial and temporal variations, we built a mixed-effect linear model[18] for the space-time mRNA profiles, which extends harmonic regression to include a spatial covariate (Figure 2B). In this model, rhythms are parameterized with cosine and sine functions, while spatial profiles are represented with (up to second order) polynomials. In its most complex form, the model uses nine parameters describing spatially modulated oscillations, and one intercept per mouse (Methods). When some of the parameters are zero, the model reduces to simpler mRNA profiles, for example purely spatial or purely temporal expression profiles (Figure 2C). We then used model selection[19] to identify the optimal parameterization and category for each gene (Methods). *In fine*, we classified each mRNA profile into one of five types of patterns (Figure 2C). If only the intercept is used, the profile will be classified as flat (F). If only time-independent zonation parameters are retained, the predicted profile will be purely zonated (Z). If only layer-independent rhythmic parameters are retained, the predicted profile will be purely rhythmic (R). If only layer-independent rhythmic parameters and time-independent zonation parameters are retained, the profile is classified as independent rhythmic-zonated (Z+R). If at least one layer-dependent rhythmic parameter is selected, the profile will be termed interacting (ZxR). This classification revealed that, overall, about 30% of the mRNA profiles were zonated (Z, Z+R and ZxR) and about 20% were rhythmic (R, Z+R and ZxR) (Figure 2D). The peak times of these rhythmic transcripts were highly consistent with bulk chronobiology data[20]

(Supplementary Figure 2B). This analysis is available as a web-app resource along with the corresponding data (https://czviz.epfl.ch).

Interestingly, we found that 7% of the analyzed genes in the liver were both zonated and rhythmic. Such dually regulated transcripts represent 25% of all zonated transcripts, and 36% of all rhythmic transcripts, respectively. The previously shown *Elovl3* transcript, involved in fatty acid elongation, and *Pck1*, a rate limiting enzyme in gluconeogenesis, are prototypical Z+R genes (Figure 1H). Gluconeogenesis requires acetyl-CoA produced via β-oxidation. As mice are in a metabolically fasted state towards the end of the light phase (~ ZT10) and oxygen needed for β -oxidation is most abundant portally[21] this process indeed needs to be both spatially and temporally regulated. Similarly, fatty acids production occurs during periods of excess energy and glycolysis (~ ZT18) and is located around the central vein[22]. The dual regulation of these genes may therefore ensure optimal liver function under switching metabolic conditions.

Dually regulated genes were mostly Z+R, with only a minority of ZxR patterns. The average expression across categories showed that rhythmic genes are more lowly expressed on average than genes in the other categories, likely reflecting shorter half-lives (Figure 2E and Supplementary Figure 2C). Together, we found that mRNA expression of many zonated genes in hepatocytes is not static, and is in fact compartmentalized both in space and time.

Properties of dually zonated and rhythmic mRNA profiles

The majority of dually regulated genes are Z+R, which denotes additive (in log) space-time effects, or dynamic patterns where slopes or shapes of spatial patterns do not change with time (Figure 2C). On the other hand, fully dynamic patterns (ZxR) are rare. Comparing the proportions of central, mid-lobular (peaking in the middle of the porto-central axis) and portal genes among the purely zonated genes (Z), and independently zonated and rhythmic genes (Z+R), did not reveal significant differences (Figure 3A), suggesting that rhythmicity is uncoupled with the direction of zonation. Similarly, comparing the phase distribution among the purely rhythmic genes (R) and the Z+R genes did not show a significant difference (Figure 3B), indicating that zonation does not bias peak expression time.

Moreover, oscillatory amplitudes were uncorrelated with the zonation slopes in Z+R genes (Figure 3C). However, we observed that large slopes (>1 or < -1, corresponding to fold changes of at least two between central and portal layers) are only associated with medium and large temporal amplitudes (>0.2, corresponding to >0.4 log2 fold change in time), as illustrated by *Elovl3*, *Rnase4* or *Pck1* (Supplementary Figure 3A). Conversely, many genes show small slopes and large amplitudes, as illustrated by *Hsp90aa1*, *Thrsp*, *Lpin1* (Supplementary Figure 3A).

Finally, for ZxR genes with potentially more complex space-time patterns, we investigated whether the spreads in amplitudes and peak times across the layers were linked (Figure 3D). This revealed that all ZxR profiles showed small spreads in peak times (<1h) but larger amplitude spreads (>0.1, log2). This phenomenon, illustrated by *Sds*, *Mt1* and *Cyp7a1* (Supplementary Figure 3B), indicates that amplitude modulation is the main factor for classification as ZxR.

smFISH analysis of space-time mRNA counts

To substantiate the RNA-seq profiles, we performed single RNA molecule fluorescence *in situ* hybridization (smFISH) experiments on a set of selected candidate genes with diverse spatio-temporal patterns. smFISH provides a sensitive, albeit low-throughput measurement of mRNA expression. We selected the core-clock genes *Bmal1* and *Per1*, classified as R in the RNA-seq analysis, *Elovl3* for Z+R, and *Acly* for the ZxR categories. Purely zonated genes (Z) were already well studied with smFISH[7]. The reconstructed scRNA-seq and smFISH profiles were consistent, though with minor discrepancies. *Bmal1* (~ZT0) and *Per1* (~ZT12) phases were nearly identical in both experiments, and the rhythms did not depend on the lobular position (Figure 4A). *Elovl3* is both centrally biased and rhythmic in RNA-seq and smFISH, even though the amplitude of the oscillations appeared lost on the portal side in the FISH experiment, presumably due to low expression and thus low signal-to-noise (Figure 4B). Finally, *Acly* showed a pattern in smFISH data which validates its classification as ZxR, especially since the amplitude is lower on the portal side where the transcript is more highly expressed (Figure 4C).

Space-time logic of hepatic functions and activity of signaling pathways

We next used our classification to understand the spatio-temporal dynamics of hepatic functions and signaling pathways in the liver. Given the prevalence of zonated gene expression profiles, we first analyzed if the circadian clock is sensitive to zonation. We found that profiles of reference core-clock genes (Supplementary Figure 4) were assigned to the rhythmic only category (R), except for *Cry1* and *Clock* that were assigned to Z+R, but with high probabilities also for R (Supplementary Tables S1 and S4). This suggests that the circadian clock is largely non-zonated and therefore robust to the heterogenous hepatic microenvironment.

We then systematically explored enrichment of biological functions in the R, Z, and Z+R categories using DAVID[23] (Supplementary Table S2). Gene clusters related to the circadian clock were clearly enriched in the R category, however, no other functions stood out as purely rhythmic. Functions of zonated genes have been described previously [7,24], here, our analysis of Z only genes highlighted that processes related to protein secretion were highly enriched in portal genes, constituents of ribosomes were strongly biased centrally, as were many P450 enzymes involved in oxidation of steroids, fatty acids and xenobiotics. Among the dually regulated Z+R genes, lipid metabolism stood out as the most enriched function, with about two thirds of central and one third of portal gene expression, showing peak times that were mostly between ZT9 and ZT21 (Figure 5A). Moreover, peroxisome related genes peaked in a similar interval with a marked preference for central patterns (Figure 5B). Lastly, genes linked with protein glycosylation peaked throughout the day and showed more portal patterns, presumably linked to portally biased protein secretion (Figure 5C). A similar analysis of enrichment in KEGG pathways, using a more exhaustive collection of backgrounds sets (Table S3) confirmed that rhythmic gene expression is spatially flat for core clock function, but zonated (Z+R) for other hepatic functions.

Next, we investigated genes targeted by reference signaling pathways previously implicated in zonation[7]. Specifically, we considered genes targeted by the Wnt, Ras and hypoxia pathways. As expected, the three pathways where strongly enriched in zonated genes (Supplementary Figure 5, column B5). In agreement with ref[7], both the positive targets of Wnt and the negative targets of Ras are strongly enriched in central genes, while the negative and positive targets, respectively, are strongly enriched in portal genes

(Supplementary Figure 5, column B4). However, no such bias was found for hypoxia. Regarding day vs. night expression, the Wnt and hypoxia targets displayed strong bias, while Ras did not (Supplementary Figure 5, column B1). Moreover, Wnt targets (both positive and negative) as well as positive targets of hypoxia showed rhythms preferentially peaking during the day (Supplementary Figure 5, column B2). Among the Z+R genes, the enrichment patterns of central (noted Z^C+R) and portal (Z^P+R) genes were similar during night and day (Supplementary Figure 5, column B3). On a finer temporal scale (Figure 6A), the rhythmic targets (R and Z+R) of Wnt and hypoxia showed a common pattern of temporal compartmentalization: the negative targets tended to be enriched around ZT0 (dark-light transition) and underrepresented around ZT14, while the positive targets were enriched around ZT10 and underrepresented around ZT20 (until about ZT3 for Hypoxia). Ras targets, positive or negative, did not exhibit significant temporal bias.

Finally, we asked whether the temporal oscillations in the expression of Wnt-activated genes might correlate with temporal oscillations in the Wnt morphogens produced by pericentral liver endothelial cells (LECs). To this end, we performed smFISH experiments and quantified the expression of the Wnt ligand *Wnt2*, as well as the Wnt antagonist *Dkk3* (Figure 6B). Consistent with the enrichment of peak times in positive Wnt targets (Figure 6A), we found that *Wnt2* expression in LECs exhibited non-uniform expression around the clock, with a significant trough at ZT18 ($p=0.0007$, Kruskal-Wallis). Although differences in *Dkk3* expression were not significant ($p=0.3$), the observed median expression was lowest at ZT12, when the expression of pericentral Wnt-targets is the highest.

Discussion

Recent genome-wide analyses of zonated gene expression in mouse and human liver[7,25] uncovered a rich organization of liver functions in space at the sublobular scale, while chronobiology studies of bulk liver tissue revealed a complex landscape of rhythmic regulatory layers orchestrated by a circadian clock interacting with feeding-fasting cycles and systemic signals[15,26–28]. Here, we established how these two regulatory programs intertwine to shape the daily space-time dynamics of gene expression patterns and

physiology in adult liver by extending our previous scRNA-seq approach[7]. We found that liver function uses gene expression programs where many genes exhibit compartmentalization in both space and time.

We chose to focus on the parenchymal cells in the liver, the hepatocytes, for which smFISH data on landmark zonated genes was readily available, which enabled reconstructing spatio-temporal mRNA profiles from scRNA-seq[7]. Our approach may be extended to other cell types in the liver; in fact, static zonation mRNA expression profiles have been obtained for LEC, using a paired-cell approach[16]. In addition, *ab initio* reconstruction methods such as diffusion pseudo time[25] or novoSpaRc[29] could be used for spatially sparse cell types with no available zonated marker genes, e.g. stellate or resident immune Kupffer cells.

To study whether the observed space-time expression profiles may be regulated by either liver zonation, 24h rhythms in liver physiology, or both, we developed a mixed-effect model, combined with model selection. This enabled classifying gene profiles into five categories representing different modes of spatio-temporal regulation, from flat to wave-like. To validate these, we performed smFISH in intact liver tissue, which showed largely compatible profiles although some quantitative differences were observed. These differences most likely reflect the limited sensitivity of RNA-seq, uncertainties in the spatial analysis of smFISH in tissues, as well as known inter-animal variability in the physiologic states of individual livers, notably related to the animal-specific feeding patterns[20].

Together, this temporal analysis confirms that a large proportion of gene expression in hepatocytes is zonated[7] or rhythmic[30], and in addition reveals marked spatio-temporal regulation of mRNA levels in mouse liver (Z+R and ZxR genes, comprising 7% of all detected genes according to our criteria). This means that zonated gene expression patterns can be temporally modulated on a circadian scale, or equivalently, that rhythmic gene expression patterns can exhibit sub-lobular structure. The dominant mode for dually regulated gene was Z+R, which corresponds to additive effects of space and time in log, or multiplicate effects in the natural coordinates, and describes genes expression patterns that are compartmentalized in both space and time. In other words, such patterns are characterized by shapes (in space) that remain invariant

with time, but whose amplitudes are rhythmically rescaled in time. Or equivalently, the oscillatory amplitude (fold change) and phases are constant along the lobular coordinate, but the mean expression is patterned along the lobule. Such multiplicative effects could reflect the combined actions of transcriptional regulators for the zone and rhythm on promoters and enhancers of Z+R genes. Indeed, gene expression changes induced by several regulators combine multiplicatively[17]. The non-zonated expression of the core clock genes we identified is compatible with such a uniform non-zonated multiplicative effect. Note that though the (relative) shape of Z+R patterns is invariant in time, threshold-dependent responses that would lie downstream of such genes would then acquire domain boundaries which can shift in time. Finally, space-time waves of gene expression (ZxR) were also observed, but to a much lesser extent, and usually with larger amplitude than phase modulation along the lobular layers.

In addition to previously discussed zonated functions[7], pathway analysis revealed that expression of ribosome protein genes is higher centrally, which, together with the previously noted zonation of proteasome components[7] could indicate that protein turnover is higher in the pericentral lobule layers. This higher turnover could preserve protein function in the stressed low-oxygen and xenobiotics-rich pericentral microenvironment. *Fatty acid metabolism* appears complex[31]: Z+R genes involved in fatty acid oxidation/degradation (*Acl1*, *Acox1*, *Ehhadh*, *Hacd3*, *Hadh*, *Hadhb*) are expressed centrally and peak around ZT12-ZT18, while *fatty acid elongation* genes either central (*Elovl3*) or portal (*Fasn*, *Elovl5*), peaking at different times of day. Interestingly, we found that expression of circadian clock transcripts is robust to metabolic zonation and is the only function showing an over-representation of rhythmic but non-zonated genes.

It was previously shown that Wnt signaling can explain the zonation of up to a third of the zonated mRNAs[7]. Wnt ligands are secreted by pericentral LECs[4,5] and form a graded spatial morphogenetic field. As a result, and as observed in our enrichment analysis (Supplementary Figure 5), Wnt-activated genes were pericentrally-zonated, whereas Wnt-repressed genes periportally-zonated. Our smFISH analysis suggested that temporal fluctuations in the expression of key ligands by pericentral LECs might account for oscillatory and zonated hepatocyte gene expression.

In summary, we demonstrate how liver gene expression can be quantitatively investigated with spatial and temporal resolution and how liver function is compartmentalized along these two axes. In particular our data suggest two scenarios: multiplicative effects of spatial and temporal regulators, and temporal regulation of spatial regulators such as WNTs.

Data availability

GEO

All data is deposited in GEO with accession code GSE145197 (reviewer token xxxx).

Web-application

The whole dataset along with the analysis is available online as a web-application at the URL https://czviz.epfl.ch/. The application was built in Python using the library *Dash by Plotly* (version 1.0).

Material and methods

Animals and ethics statement

All animal care and handling were approved by the Institutional Animal Care and Use Committee of WIS and by the Canton de Vaud laws for animal protection (authorization VD3197.b). Male C57bl6 mice aged of 6 weeks, housed under reverse-phase cycle and under ad libitum feeding were used to generate sc-RNA-seq data of hepatocytes and single-molecule RNA-FISH (smFISH). Male mice between 8 to 10 weeks old, housed under 12:12 light-dark cycle, and having access to food only during the night (restricted-feeding) were used for smFISH of circadian clock genes.

Hepatocytes isolation and single-cell RNA-seq

Liver cells were isolated using a modified version of the two-step collagenase perfusion method of Seglen[31]. The tissue was digested with Liberase Blendzyme 3 recombinant collagenase (Roche Diagnostics) according to the manufacturer instructions. To enrich for hepatocytes, we applied a centrifuge step at 30g for 3 min to pull down all hepatocytes while discarding most of the non-parenchymal cells that remained in the sup. We next enriched for live hepatocytes by percoll gradient, hepatocytes pellet was

resuspend in 25 ml of PBS, percoll was added for a final concentration of 45% and mixed with the hepatocytes. Dead cells were discarded after a centrifuge step (70g for 10min) cells were resuspended in 10x cells buffer (1x PBS, 0.04% BSA) and proceeded directly to the 10x pipeline. The cDNA library was prepared with the 10X genomics Chromium system according to manufactures instructions and sequencing was done with Illumina Nextseq 500.

Filtering of raw scRNA-seq data

The initial data analysis was done in R v3.4.2 using Seurat v2.1.0[32]. Each expression matrix was filtered separately to remove dead, dying and low quality cells. We firstly only kept genes that were expressed in at least 5 cells in any of the ten samples. We then defined a set of valid cells with more than 500 expressed genes and between 1000 and 10000 unique molecular identifiers (UMIs) and secondly an additional expression matrix with cells having between 100 and 300 UMIs which was used for background estimation. Other UMI-filters have been tried, but yielded equal or less reliable profiles. The mean expression of each gene was then calculated for the background dataset and subtracted from the set of valid cells. This was subsequently filtered to only include hepatocytes by removing cells with expression of non-parenchymal liver cell genes. Next, the cells were filtered based on the fraction of mitochondrial gene expression. First, expression levels in each cell were normalized by the sum of non-mitochondrial and non-major urinary protein (*Mup*) genes. Indeed, as mitochondria are more abundant in periportal hepatocytes, the expression of mitochondrial genes is higher in this area[33]; and since these genes are very highly expressed, including them would reduce the relative expression of all other genes based on the cell's lobular location. *Mup* genes are also highly abundant and mapping their reads to a reference sequence is unreliable due to their high sequence homology[34]. Then, to assess the quality of the selected sets of cells, we computed the anti-correlation between the expression levels of *Cyp2e1* and *Cyp2f2*, two strongly and oppositely zonated genes, across all cells. Cells with 9-35% mitochondrial gene expression yielded the best quality, and we used these as input for the spatial reconstruction algorithm.

Spatial reconstruction of zonation profiles from scRNA-seq data

Choice of landmark genes.

The reconstruction algorithm relies on *a priori* knowledge about the zonation of a small set of landmark genes to infer the location of the cells. Ref[7] used smFISH to determine the zonation pattern *in situ* of 6 such landmark genes and used them to reconstruct the spatial profiles of all other genes at a single time point. Since we here aimed at reconstructing zonation profiles at different time points, we could not rely on those landmark genes, which might be subject to temporal regulation. Therefore, we used an alternative strategy where we selected landmark zonated genes from Ref[7] ($q < 0.2$), with the additional constraints that those should be highly expressed (mean expression in fraction UMI of more than 0.01% and less than 0.1%), and importantly vary little across mice and time. Specifically, we calculated the variability in the mean expression (across all layers) between all mice for every gene and removed genes with >= 10% variability. This yielded 27 central (*Akr1c6, Alad, Blvrb, C6, Car3, Ccdc107, Cml2, Cyp2c68, Cyp2d9, Cyp3a11, Entpd5, Fmo1, Gsta3, Gstm1, Gstm6, Gstt1, Hpd, Hsd17b10, Inmt, Iqgap2, Mgst1, Nrn1, Pex11a, Pon1, Psmd4, Slc22a1, Tex264*); and 28 portal (*Afm, Aldh1l1, Asl, Ass1, Atp5a1, Atp5g1, C8a, C8b, Ces3b, Cyp2f2, Elovl2, Fads1, Fbp1, Ftcd, Gm2a, Hpx, Hsd17b13, Ifitm3, Igf1, Igfals, Khk, Mug2, Pygl, Sepp1, Serpina1c, Serpina1e, Serpind1, Vtn*) landmark genes.

Reconstruction algorithm.

The reconstruction algorithm is based on the algorithm in Ref[7] and was used in the modified version from Ref[16]. The procedure was applied independently on each mouse, yielding ten spatial gene expression profiles for each gene, given as fraction of UMI per cell.

Spatiotemporal analysis of liver gene expression profiles

Data

Each profile for the 14678 genes includes 8 layers from the pericentral to the periportal zone and 4 time points: ZT0 (n=3 biological replicates from individual mice), ZT6 (n=2),

ZT12 (n=3) and ZT18 (n=2). The expression levels (noted as x) are then log-transformed as follows:

$$y = log_2(x + \Delta) - B \quad (i)$$

The offset $\Delta = 10^{-4}$ buffers variability in lowly expressed genes, while the shift $B = -log_2(11 \times 10^{-5})$ changes the scale so that $y = 0$ corresponds to about 10 mRNA copies per cell (we expect on the order of 1M mRNA transcripts per liver cell).

Reference genes

For ease of interpretation (Figure 2 and Supplementary Figure 2), we used a set of reference circadian genes and a set of reference zonated genes, highlighted in several figures.

The reference core circadian clock and clock output genes are the following: *Bmal1, Clock, Npas2, Nr1d1, Nr1d2, Per1, Per2, Cry1, Cry2, Dbp, Tef, Hlf, Elovl3, Rora, Rorc*.

The reference zonated genes are the following: *Glul, Ass1, Asl, Cyp2f2, Cyp1a2, Pck1, Cyp2e1, Cdh2, Cdh1, Cyp7a1, Acly, Alb, Oat, Aldob, Cps1*.

Gene expression variance in space and time

To analyze variability in space and time (Figure 2A) we computed, for each gene, the spatial variance V_x and the temporal variance V_T. Let $y_{x,t,j}$ represent the expression profile, with j the replicate index, $t \in \{1,2,\dots,N_t\}$ the time index, and $x \in \{1,2,\dots,N_x\}$ the layer index. Then, V_x and V_T are computed as follows:

$$V_X = \frac{1}{N_t}\sum_t \frac{\sum_x [\sum_j (y_{x,t,j} - \frac{1}{N_x}\sum_x y_{x,t,j})]^2}{N_j^2 N_x} \quad (ii)$$

$$V_T = \frac{1}{N_x}\sum_x \frac{\sum_t [\sum_j (y_{x,t,j} - \frac{1}{N_t}\sum_t y_{x,t,j})]^2}{N_j^2 N_t} \quad (iii)$$

Thus, the spatial variance V_x is computed along the space (and averaged over the replicates) for each time condition, and then averaged over time. The procedure is similar, symmetrically, for V_t.

Genes filtering

For the analyses in Figure 2, we selected transcripts that were reproducible between replicates, as well as sufficiently highly expressed (see scatterplot in Supplementary Figure 2A). To assess reproducibility across replicates, we computed the average relative variance of the spatiotemporal profiles over the replicates:

$$V_J = \frac{\frac{1}{N_x N_t}\sum_{x,t}\left[\frac{1}{N_j}\sum_j\left(y_{x,t,j} - \frac{1}{N_j}\sum_j y_{x,t,j}\right)^2\right]}{\frac{1}{N_x N_t N_j}\sum_{x,t,j}\left[(y_{x,t,j} - \frac{1}{N_x N_t N_j}\sum_{x,t,j} y_{x,t,j})^2\right]} \quad (iv)$$

We considered genes with values below 50% (Supplementary Figure 2). To filter lowly expressed genes, we required the maximum expression level across layers and time points (fraction of UMIs) to exceed 10^{-5} which corresponds to y = 0 or about 10 copies of mRNA per cell. While this kept most of the reference zonated and circadian genes, the filtering selected a total of 5085 genes, used for all analyses presented in Figures 2-5. In addition, we systematically discarded *Cyp2a4* and *Cyp2a5* from analyses, as these genes are not discernable from scRNA-seq due to their highly similar sequences.

Mixed-effect model for spatiotemporal mRNA profiles

Since the data is longitudinal is space (8 layers measured in each animal), modelling the space-time profiles require the use of mixed-effect models. To systematically analyze the spatiotemporal mRNA profiles, we used a parameterized function. Specifically, the model uses sine and cosine functions for the time, and polynomials (up to degree 2) for space. Possible interaction between space and time are described as space-dependent oscillatory functions, or equivalently, time-dependent polynomial parameters. Our model for the transformed mRNA expression *y* reads:

$$y_{x,t,i} = \mu_i + \mu(x) + a(x)\cos(\omega t) + b(x)\sin(\omega t) + \varepsilon_{x,t,i} \qquad (v)$$

Here t is the time, x the spatial position along the liver layers, and $i \in \{1,2,\dots,10\}$ the animal index. This function naturally generalizes harmonic regression, often used for analysis of circadian gene expression[20], by introducing space-dependent coefficients:

$$\begin{cases} \mu(x) &= \mu_0 + \mu_1 P_1(x) + \mu_2 P_2(x) \\ a(x) &= a_0 + a_1 P_1(x) + a_2 P_2(x) \\ b(x) &= b_0 + b_1 P_1(x) + b_2 P_2(x) \end{cases}$$

Here, P_1 and P_2 are the Legendre polynomials of degrees 1 and 2, respectively; μ_0, μ_1 and μ_2 represent the static zonation profile, a_0 and b_0 represent the global (space-independent) rhythmicity of the gene, while a_1, a_2, b_1, b_2 represent layer-dependent rhythmicity. $\varepsilon_{x,t,j}$ is a Gaussian noise term with standard deviation σ. In addition to the fixed-effect parameters described so far, we also introduced a mouse-specific random-effect μ_i (with zero mean). This parameter groups the dependent layer measurements (obtained in the same animal) and thereby properly adjusts the biological sample size for the rhythmicity analysis.

Phases φ (related peak times t through $t = \varphi * 24/2\pi$) and amplitudes A, for each profile can then be computed for any layer from the coefficients $a(x)$ and $b(x)$:

$$\varphi(x) = arctan2\big(b(x), a(x)\big) \qquad A(x) = \sqrt{a(x)^2 + b(x)^2} \qquad (vi)$$

We also note that an equivalent writing of the model formulates the problem in terms of time-dependent zonation parameters instead of space-dependent rhythmicity:

$$y_{x,t,i} = \mu_i + \mu_0(t) + \mu_1(t)P_1(x) + \mu_2(t)P_2(x) + \varepsilon_{x,t,i} \qquad (vii)$$

where:

$$\begin{cases} \mu_0(t) & = & \mu_0 + a_0\cos(\omega t) + b_0\sin(\omega t) \\ \mu_1(t) & = & \mu_1 + a_1\cos(\omega t) + b_1\sin(\omega t) \\ \mu_2(t) & = & \mu_2 + a_2\cos(\omega t) + b_2\sin(\omega t) \end{cases}$$

In this study, we fixed $\omega = \frac{2\pi}{24\,h}$ since the animals were entrained in a 24 h light-dark cycle and the low time resolution would prevent us from studying ultradian rhythms.

The model parameters, including the variance of the random effects and Gaussian noise strength σ, are estimated for each gene using the *fit* function from the Python library StatsModels (version 0.9.0). Nelder-Mead was chosen as the optimization method, and the use of a standard likelihood was favored over the REML likelihood to allow for model comparison[35]. To prevent overfitting of the gene profiles, we added a noise offset $\sigma_0 = 0.15$ [log_2] to the estimated noise σ, in the expression of the likelihood function used in the mixed-effect model optimization.

Depending on the gene, the model presented in (v) and (vii) may be simplified by setting all or some of the (fixed) parameters to 0. For example, a non-oscillatory gene profile would normally have non-significant a_j and b_j parameters. In practice,

considering the fixed effects, 2^9 sub-models of various complexity can be generated. However, we added a few reasonable requirements to reduce the number of models. First, the intercept μ_0 must be present in every model. Similarly, the parameters a_0 and b_0, providing a global rhythm, must be present in every rhythmic model. Finally, the parameters a_j and b_j for *j*=0,1,2 must be paired to ensure a proper phase definition (*vi*).

The models can then be classified in different categories, depending on the retained (non-zero) parameters (Figure 2C):

- The model comprising only the intercepts μ_0 and μ_i, termed flat (F).
- The models comprising only the intercepts and zonation parameters: $\mu_1$and/or μ_2, termed purely zonated (Z).
- The models comprising only the intercepts and rhythmic parameters: a_0 and b_0, termed purely rhythmic (R).
- The models comprising only the intercepts, zonated parameters and rhythmic parameters: μ_1 and/or μ_2, and a_0, b_0, termed independent (Z+R).
- The models comprising interaction parameters: a_j and b_j for *j*=1,2, termed interacting (ZxR).

Note that, since the random-effect parameter μ_i is interpreted as a correction for the data rather than a predictive parameter, it is systematically used for the fits, but not plotted in the final retained profiles (e.g. Figures 1, Supplementary Figure 1).

The Bayesian Information Criterion (BIC) is then used for model selection, enabling to choose the best model for a given gene profile. It appears that, for some profiles, several competing models can result in very close BIC values (see e.g. the discussion on *Clock* and *Cry*1 in the Results). Therefore, if some models have a relative difference of less than 1% in their BIC (sorted by increasing order), we systematically keep the one with the highest number of parameters. We assign probabilities to the different categories (F, Z, R, Z+R and ZxR), computed as Schwartz BIC weights[36] (Supplementary Table S4). All best fits with their parameter values are listed in Supplementary Table 1.

Comparison of peak times with the dataset from Atger et al.

We compared our rhythmically classified genes with those obtained from the data in[20]. These data consist of bulk liver RNA-seq data sampled every 2 hours for 24 hours, with 4

replicates per time condition. Thus, there are two main differences between our current and that dataset. First, we can only compare the genes for which rhythmicity is not changing across layers, *viz.* the R and Z+R categories. Secondly, our dataset has a lower temporal resolution, with fewer replicates per time point, meaning that the statistical power for the detection of rhythmic genes is higher in ref[20]. We therefore did not compare genes found as flat in our dataset.

To assess gene rhythmicity from ref[20], we used harmonic regression on the log-transformed profiles as previously. Using the same notation as above, we define the two following models:

$$\begin{cases} y_{t,i} &= \mu + \varepsilon & (ix) \\ y_{t,i} &= \mu + a\cos(\omega t) + b\sin(\omega t) + \varepsilon & (x) \end{cases}$$

We then fit eq. $(viii)$ and eq. (ix) to every transcript, and, for each of them, keep the model with the lowest BIC. Transcripts for which eq. (ix) is favored are assigned to the flat category (F), while transcripts for which eq. (x) is favored are assigned to the rhythmic category (R). We then compare the phases and amplitudes of the transcripts classified as rhythmic in both datasets, and compute the circular correlation coefficient[37].

KEGG and DAVID pathway Enrichment analysis

Functional annotation clustering from DAVID[23] for the three categories Z, R, Z+R was ran with standard parameters, using the above set of 5002 selected genes as background set.

We also studied the enrichment of KEGG pathways using hypergeometric tests. In practice, if a background set contains N genes, among which K belong to the category Y, and if a given pathway contains n genes, then the p-value corresponding to an enrichment of k genes is the probability of observing at least k genes from the pathway which belong to the category Y, that is (for a right-sided test):

$$p(x \geq k) = \sum_{x=k}^{n} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \quad (xi)$$

FDR is then adjusted running the Benjamini-Hochberg procedure on the resulting set of p-values, for each category (zonated, rhythmic, etc.) separately. For clarity, a value z is also computed alongside as:

$$z = \frac{k - \frac{K}{N}n}{\frac{K}{N}n} \quad (xii)$$

This value must be interpreted as the relative difference between the observed number of genes k in the pathway belonging to Y, and the expected value if the distribution was similar as the one from the background.

332 KEGG pathways were tested against KEGG genome T01002. To exclude too specific and too general pathways, pathways having less than 3 or more than 500 genes in common with our dataset were discarded from the analysis.

smFISH

Analysis of Z+R and ZxR genes (Stellaris smFIH probes)

Preparation of probe libraries, hybridization procedure and imaging conditions were previously described[16]. smFISH probe libraries were coupled to TMR, Alexa594 or Cy5. Cell membranes were stained with alexa fluor 488 conjugated phalloidin (Rhenium A12379) that was added to GLOX buffer[38]. Portal node was identified morphologically on DAPI images based on bile ductile, central vein was identified using smFISH for Glul in TMR, included in all hybridizations. For zonation profiles, images were taken as scans spanning the portal node to the central vein. Images were analyzed using ImageM[38]. Quantification of zonation profiles in different circadian time point were generated by counting dots and dividing the number of dots in radial layers spanning the porto-central axis by the layer volume. Central vein niche NPCs were identified by co-staining of Aqp1, Igfbp7 and Ptprb. The central vein area was imaged and the images were analyzed using ImageM. We counted dots of Wnt2 and Dkk3 expression in NPCs lining the central vein and removed background dots larger than 25 pixels. We then divided the dot count by the segmented cell volume. In total 489 NPCs from 120 central veins of 2 mice (ZT0,6,12,18) were imaged and a Kruskal-Wallis test based on the mean mRNA dot concentration in each cell was performed to compare the timepoints.

Temporal analysis of circadian genes (RNA scope smFISH probes)

smFISH of R genes were done on fresh-frozen liver cryosections (8μm) embedded in O.C.T Compound (Tissue-Tek; Sakura-Finetek USA), sampled every three hours (ZT0 to ZT21). RNAscope® probes for *Bma1l* mRNA (Mm-Arntl, catalog #: 438748-C3) and *Per1* mRNA (Mm-Per1, catalog #: 438751) were used, according to the manufacturer's instructions for the RNAscope Fluorescent Multiplex V1 Assay (Advanced Cell Diagnostics). To detect the central vein, an immunofluorescence of Glutamine Synthetase (ab49873, Abcam, diluted 1:2000 in PBS/BSA 0.5%/Triton-X0.01%) was done together with smFISH. Nuclei were counterstained with DAPI and sections were mounted with ProLong™ Gold Antifade Mountant. Liver sections were imaged with a Leica DM5500 widefield microscope and an oil-immersion x63 objective. Z-stacks were acquired (0.2μm between each Z position) and mRNA transcripts were quantified using ImageJ, as described previously in ref[26]. Pericentral (PC) and Periportal (PP) veins were manually detected based on Glutamine Synthetase IF or on bile ducts (DAPI staining). The Euclidean distance between two veins and the distance from the vein of each mRNA transcript were calculated. mRNA transcripts were assigned to a PP or PC zone if the distance from the corresponding vein was smaller than one-third of the distance between the PP and PC veins (ranging from 50 to 130μm).

Wnt2 and Dkk3 expression in LEC

Preparation of probe libraries, hybridization protocol and imaging conditions were previously described[16]. The *Aqp1, Igfbp7 and Ptprb* probe libraries were coupled to TMR, the *Wnt2* library was coupled to *Cy3* and the *Dkk3* library was coupled to *Cy5*. Cell membranes were stained with alexa fluor 488 coupled to phalloidin (Rhenium A12379) that was added to GLOX buffer[38]. The central vein was identified based on morphological features inspected in the DAPI and Phalloidin channels and presence of Wnt2-mRNA (detected by smFISH). Endothelial cells were identified by co-staining of *Aqp1*, *Igfbp7* and *Ptprb*. The central vein area was imaged and the images were analyzed using ImageM[38]. We counted dots of *Wnt2* and *Dkk3* expression (corresponding to single mRNA molecules) in endothelial cells lining the central vein and removed background dots larger than 25 pixel3. We then divided the dot count by the segmented cell volume. In total 516 endothelial cells from 120 central veins of at least 2 mice per time point.

Acknowledgments

This work was supported by the Rothschild Caesarea Foundation' fund managed by Weizmann Institute and EPFL, a Swiss National Science Foundation Grant 310030_173079 (to F.N.), and the EPFL. 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 Sieff of Brimpton Memorial Fund, the Wolfson Foundation SCG, the Wolfson Family Charitable Trust, Edmond de Rothschild Foundations, 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 Chan Zuckerberg Initiative grant no. CZF2019-002434, 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 program (grant agreement no. 768956, the Bert L. and N. Kuggie Vallee Foundation, and the Howard Hughes Medical Institute (HHMI) international research scholar award.

Figure captions

Figure 1. A scRNA-seq approach to space-time gene expression patterns in mouse liver.

(A-E) *t*-SNE visualizations of the scRNA-seq data (19663 hepatocytes from n = 10 mice). Individual cells are colored by the lobule layer (A), Zeitgeber time (B), expression levels of the zonated genes *Cyp2f2* and *Cyp2e1* (C-D), or the circadianly regulated gene *Elovl3* (E).

(F-H) Reconstructed spatial profiles (lobule layers 1-8) of selected zonated but temporally static genes (F, top: *Glul* pericentrally (PC) expressed, bottom: *Ass1* periportally (PP) expressed); rhythmic but non-zonated genes (G, top: *Bmal1* peaking at ZT0, bottom: *Dbp*, peaking at ZT6-12); zonated and rhythmic genes (H, Top: *Elovl3*, bottom: *Pck1*). Expression levels correspond to fraction of total UMI per cell in linear scale. Log-transformed profiles are in Supplementary Figure 1. Dots in F-H represent data points from the individual mice.

Figure 2. Space-time mRNA expression profiles categorized with mixed-effect models.

(A) Spatial and temporal variation for each mRNA transcript profile, calculated as standard deviations (SD) of log2 expression along spatial or temporal dimensions (Methods). Colored dots correspond to reference zonated genes (orange) and reference rhythmic genes (blue) (Methods).

(B) Generalized harmonic regression model for spatio-temporal expression profiles describing a static but zonated layer-dependent mean $\mu(x)$, as well as layer-dependent rhythmic amplitudes ($a(x)$ and $b(x)$). All layer-dependent coefficients are modeled as second order polynomials; i denotes the biological replicates. μ_i are random effects needed due to the asymmetric experimental design of the study (Methods).

(C) Schema illustrating the different categories of profiles. Depending on which coefficients are non-zero (Methods), genes are assigned to: F (flat), Z (purely zonated), R (purely rhythmic), Z+R (additive zonation and rhythmicity), ZxR (interacting zonation and rhythmicity). Graphs emphasize either zonation (top), with the x-axis representing layers, or rhythmicity (bottom), with the x axis representing time (ZT). Right side of the panel: an example of fit (*Elovl3*).

(D) Number of transcripts in each category.

(E) Boxplot of the mean expression per category shows that the zonated genes (Z, Z+R and ZxR) are more expressed than flat (F) and rhythmic genes (R). Complex modulated genes (ZxR) are the most expressed (median).

Figure 3. Properties of dually zonated and rhythmic mRNA profiles

(A) Proportions of pericentral (green) and periportal (blue) transcripts are similar in Z and Z+R. Mid-lobular genes (orange) are rare (<2%).

(B) Peak time distribution of rhythmic transcripts are similar in R and Z+R categories.

(C) Effect sizes of zonation (slope) *vs.* rhythmicity (amplitude) in Z+R genes.

(D) Magnitude of time shifts (delta time, in hours) *vs.* amplitude gradient (delta amplitude, in log2) along the central-portal axis in ZxR genes.

Figure 4. smFISH analysis of rhythmic and zonated transcripts.

(A): smFISH (RNAscope, Methods) of the core clock genes *Bmal1* and *Per1* (both in R category) in liver slices sampled every 3 hours. **Left:** representative images at ZT0, ZT06, ZT12 and ZT18 for *Bmal1*. Pericentral veins (CV) and a periportal node (PN) are marked. Scale bas is 50µm. Endothelial cells lining the PC and cholangiocytes surrounding the PP were excluded from the quantification. mRNA transcripts and nuclei were detected in PN and PC zones (Methods). **Right:** Temporal profiles of *Bmal1I* and *Per1* from smFISH (top, quantification of the number of mRNA transcripts at 8 time points ZT0 to ZT21, every 3 hours, shaded area indicate SD across images), in PN and PC regions, and scRNA-seq (bottom, shaded areas is SD across mice).

(B-C) smFISH (Stellaris, Methods) for *Elovl3* (Z+R) and *Acly* (ZxR). smFISH quantifications were made for ZT0 and ZT12 (Methods). Left: representative images at ZT0, ZT12 for *Elovl3* (B) or *Acly* (C). Pericentral veins (CV) and a periportal node (PN) are marked. Scale bar - 20µm. Right: quantified profiles for each gene in the two time points (top), prediction of profiles at four time points base on the scRNA-seq (bottom).

Figure 5. Space-time logic of compartmentalized hepatic functions for Z+R genes.

(A-C) Polar representations of central (left) and portal (right) genes from three prominently enriched pathways in the Z+R category (Table S2). Peak expression times are arranged clockwise (ZT0 on vertical position) and amplitudes (log2, values indicated on the radial axes) increase radially. q indicates significance of enrichment (Bonferroni adjusted).

Figure 6. Wnt-targets correlates with *Wnt2* expression.

(A) Enrichment/depletion of genes targeted by the Wnt, Ras and Hypoxia pathways with respect to the rhythmic genes (R and Z+R) peaking at different time of the day (window size: 3h). Colormap show p-values (two-tailed hypergeometric test): red (blue) areas indicate times of the day for which there are more (less) pathways genes peaking than expected.

(B) Left: Representative smFISH images of *Wnt2* and *Dkk3* expression in endothelial cells at ZT0 (top) and ZT18 (bottom), nuclei are stained in blue (DAPI) and cell membrane in grey (phalloidin). Scale bars, 20 µm; inset scale bars 5 µm; CV = central

vein. Right: Quantitative analysis of smFISH images (516 cells of 120 central veins of at least two mice per time point. mRNA expression is in smFISH dots per μm^3.

References

1. Gebhardt, R. Metabolic zonation of the liver: regulation and implications for liver function. *Pharmacol. Ther*. **53**, 275–354 (1992).
2. 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*. **107**, 10371–10376 (2010).
3. Ben-Moshe, S. & Itzkovitz, S. Spatial heterogeneity in the mammalian liver. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 395–410 (2019).
4. 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).
5. Planas-Paz, L. *et al.* The RSPO–LGR4/5–ZNRF3/RNF43 module controls liver zonation and size. *Nat. Cell Biol.* **18**, 467–479 (2016).
6. Jungermann, K. & Kietzmann, T. Zonation of parenchymal and nonparenchymal metabolism in liver. *Annu. Rev. Nutr*. **16**, 179–203 (1996).
7. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
8. Dibner, C., Schibler, U. & Albrecht, U. The Mammalian Circadian Timing System: Organization and Coordination of Central and Peripheral Clocks. *Annu. Rev. Physiol.* **72**, 517–549 (2010).
9. Green, C. B., Takahashi, J. S. & Bass, J. The Meter of Metabolism. *Cell* **134**, 728–742 (2008).
10. Yeung, J. & Naef, F. Rhythms of the Genome: Circadian Dynamics from Chromatin Topology, Tissue-Specific Gene Expression, to Behavior. *Trends Genet. TIG* **34**, 915–926 (2018).
11. Maury, E., Ramsey, K. M. & Bass, J. Circadian Rhythms and Metabolic Syndrome: From Experimental Genetics to Human Disease. *Circ. Res.* **106**, 447–462 (2010).
12. Mermet, J., Yeung, J. & Naef, F. Systems Chronobiology: Global Analysis of Gene Regulation in a 24-Hour Periodic World. *Cold Spring Harb. Perspect. Biol.* **9**, a028720 (2017).
13. Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A

circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16219–16224 (2014).

14. Sobel, J. A. *et al.* Transcriptional regulatory logic of the diurnal cycle in the mouse liver. *PLOS Biol.* **15**, e2001069 (2017).

15. Wang, J. *et al.* Nuclear Proteomics Uncovers Diurnal Regulatory Landscapes in Mouse Liver. *Cell Metab.* **25**, 102–117 (2017).

16. Halpern, K. B. *et al.* Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. *Nat. Biotechnol.* **36**, 962–970 (2018).

17. Beal, J. Biochemical complexity drives log-normal variation in genetic expression. *Eng. Biol.* **1**, 55–60 (2017).

18. McLean, R. A., Sanders, W. L. & Stroup, W. W. A Unified Approach to Mixed Linear Models. *Am. Stat.* **45**, 54 (1991).

19. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196 (2004).

20. Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc. Natl. Acad. Sci.* **112**, E6579–E6588 (2015).

21. Kietzmann, T. Metabolic zonation of the liver: The oxygen gradient revisited. *Redox Biol.* **11**, 622–630 (2017).

22. Jungermann, K. Dynamics of zonal hepatocyte heterogeneity. Perinatal development and adaptive alterations during regeneration after partial hepatectomy, starvation and diabetes. *Acta Histochem. Suppl.* **32**, 89–98 (1986).

23. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

24. Gebhardt, R. Liver zonation: Novel aspects of its regulation and its impact on homeostasis. *World J. Gastroenterol.* **20**, 8491 (2014).

25. Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019).

26. Mermet, J. *et al.* Clock-dependent chromatin topology modulates circadian transcription and behavior. *Genes Dev.* **32**, 347–358 (2018).

27. Le Martelot, G. *et al*. Genome-Wide RNA Polymerase II Profiles and RNA Accumulation Reveal Kinetics of Transcription and Associated Epigenetic Changes During Diurnal Cycles. *PLoS Biol.* **10**, e1001442 (2012).

28. Wang, J. *et al.* Circadian clock-dependent and -independent posttranscriptional regulation underlies temporal mRNA accumulation in mouse liver. *Proc. Natl. Acad. Sci.* **115**, E1916–E1925 (2018).

29. Nitzan, M., Karaiskos, N., Friedman, N. & Rajewsky, N. Gene expression cartography. *Nature* (2019) doi:10.1038/s41586-019-1773-3.

30. Vollmers, C. *et al*. Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21453–21458 (2009).

31. Seglen, P. O. Chapter 4 Preparation of Isolated Rat Liver Cells. in *Methods in Cell Biology* vol. 13 29–83 (Elsevier, 1976).

32. 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).

33. Loud, A. V. A QUANTITATIVE STEREOLOGICAL DESCRIPTION OF THE ULTRASTRUCTURE OF NORMAL RAT LIVER PARENCHYMAL CELLS. *J. Cell Biol*. **37**, 27–46 (1968).

34. Kuhn, N. J., Woodworth-Gutai, M., Gross, K. W. & Held, W. A. Subfamilies of the mouse major urinary protein (MUP) multi-gene family: sequence analysis of cDNA clones and differential regulation in the liver. *Nucleic Acids Res.* **12**, 6073–6090 (1984).

35. Müller, S., Scealy, J. L. & Welsh, A. H. Model Selection in Linear Mixed Models. *Stat. Sci.* **28**, 135–167 (2013).

36. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **6**, 461–464 (1978).

37. Jammalamadaka, S. R. & SenGupta, A. *Topics in Circular Statistics*. vol. 5 (WORLD SCIENTIFIC, 2001).

38. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in mammalian tissue. *Nat. Protoc.* **8**, 1743–1758 (2013).

Figure 1

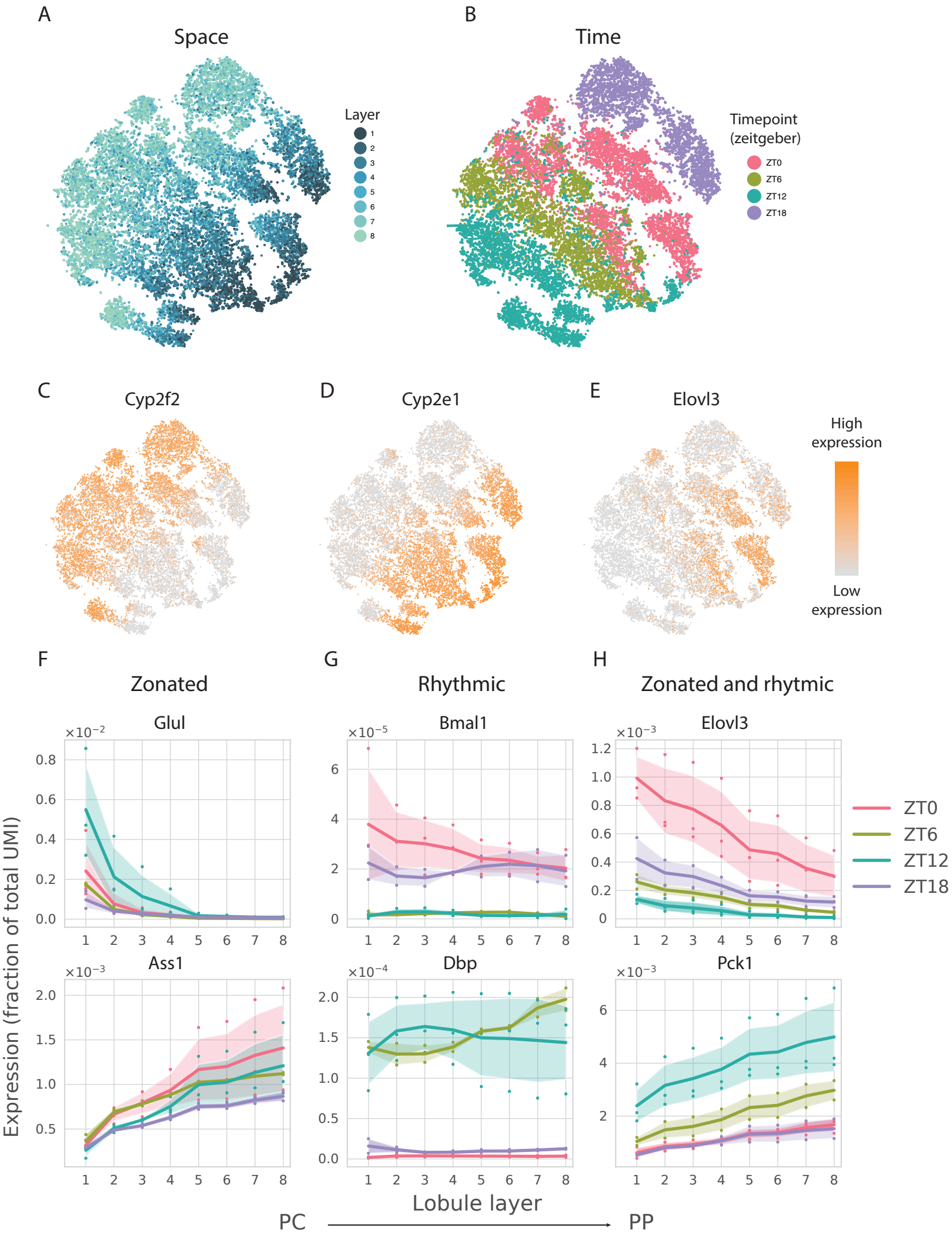
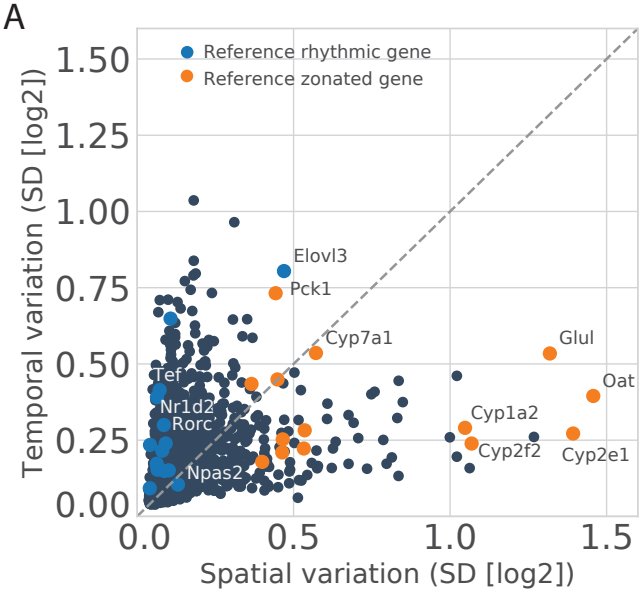

A
Space
Layer
1
2
3
4
5
6
7
8
B
Time
Timepoint
(zeitgeber)
ZT0
ZT6
ZT12
ZT18
C
Cyp2f2
D
Cyp2e1
E
Elovl3
High
expression
Low
expression
F
Zonated
G
Rhythmic
H
Zonated and rhytmic
Glul
Bmal1
Elovl3
Ass1
Dbp
Pck1
Expression (fraction of total UMI)
Lobule layer
PC
PP

Figure 2

A

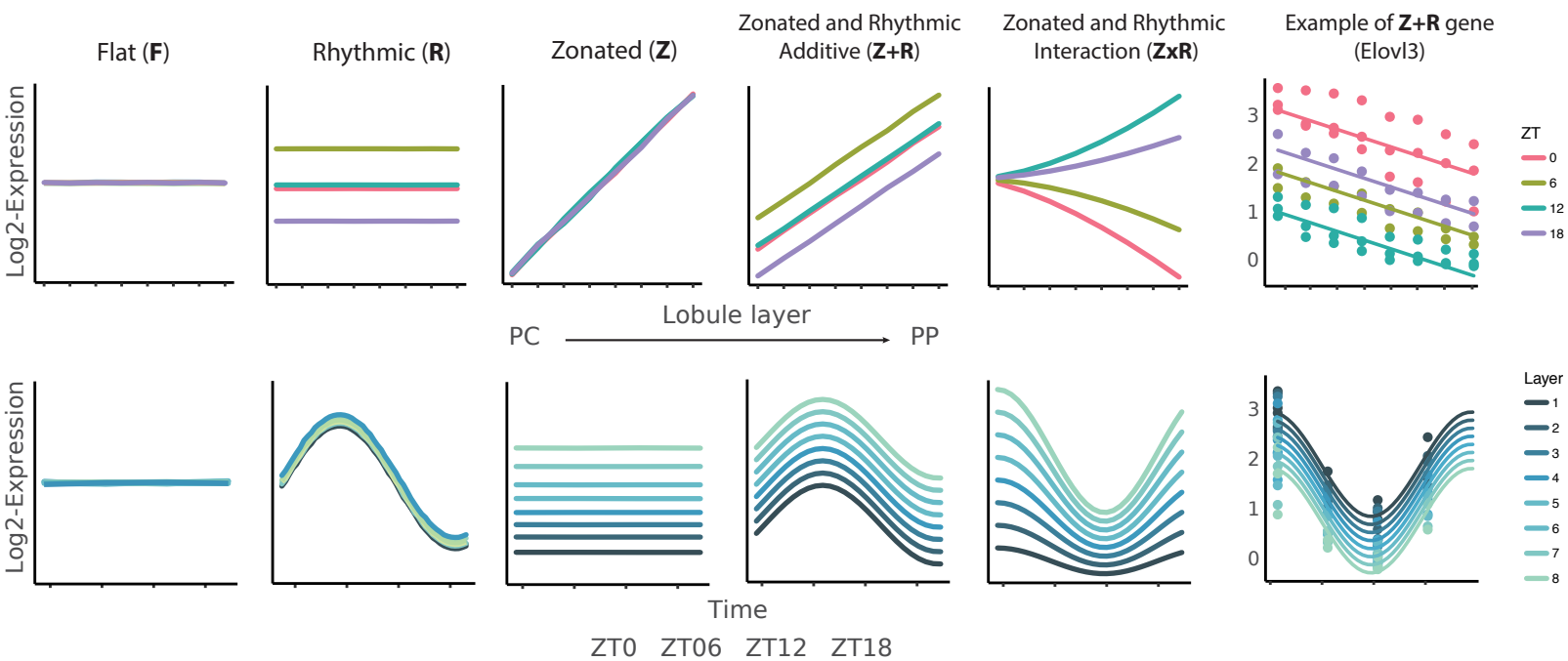


B

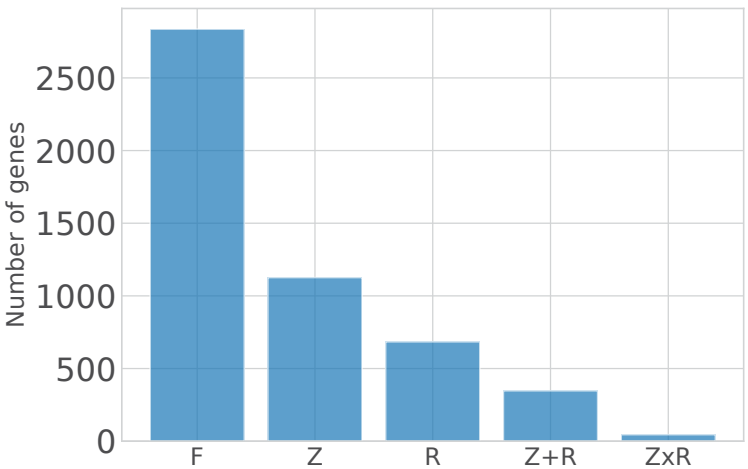
$$y_{x,t,i} = \mu_i \text{ : animal-specific random-effect}$$
$$+ \mu(x) \text{ : static zonation}$$
$$\left.\begin{array}{l} + a(x)\cos(\omega t) \\ + b(x)\sin(\omega t) \end{array}\right\} \text{: zone-dependent rhythmicity}$$
$$+ \epsilon_{x,t,i} \text{ : noise}$$

With:
$$\mu,\ a,\ b\text{: polynomials}$$
$$\omega = \frac{2\pi}{24}\text{: angular frequency } (\mathrm{h}^{-1})$$

C



D



E

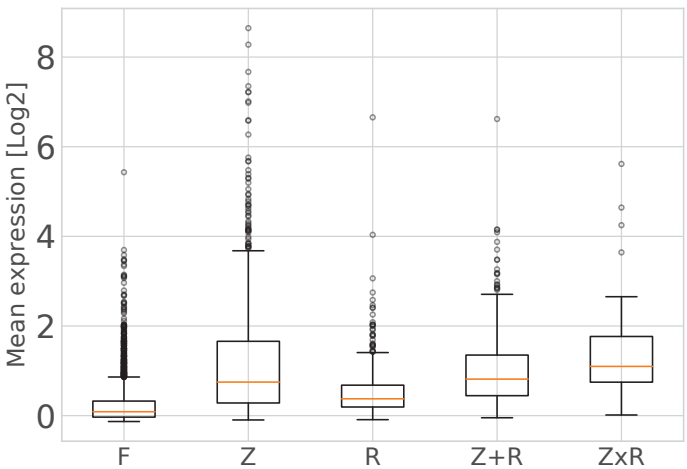


Figure 3

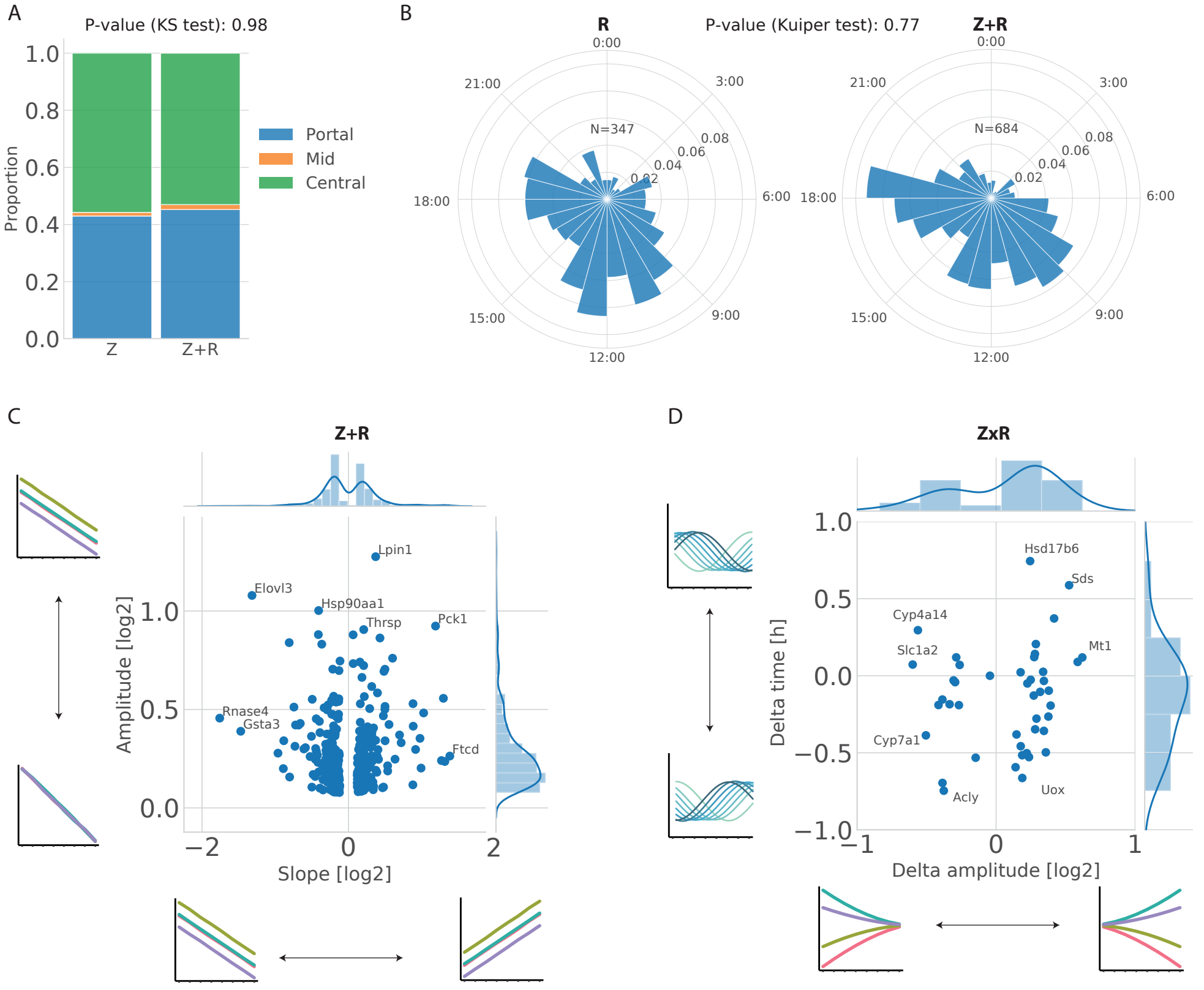

A
P-value (KS test): 0.98
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Z
Z+R
Portal
Mid
Central
B
R
P-value (Kuiper test): 0.77
Z+R
N=347
N=684
0:00
3:00
6:00
9:00
12:00
15:00
18:00
21:00
0.02
0.04
0.06
0.08
C
Z+R
Amplitude [log2]
Slope [log2]
Lpin1
Elovl3
Hsp90aa1
Thrsp
Pck1
Rnase4
Gsta3
Ftcd
D
ZxR
Delta time [h]
Delta amplitude [log2]
Hsd17b6
Sds
Cyp4a14
Slc1a2
Mt1
Cyp7a1
Acly
Uox

Figure 4

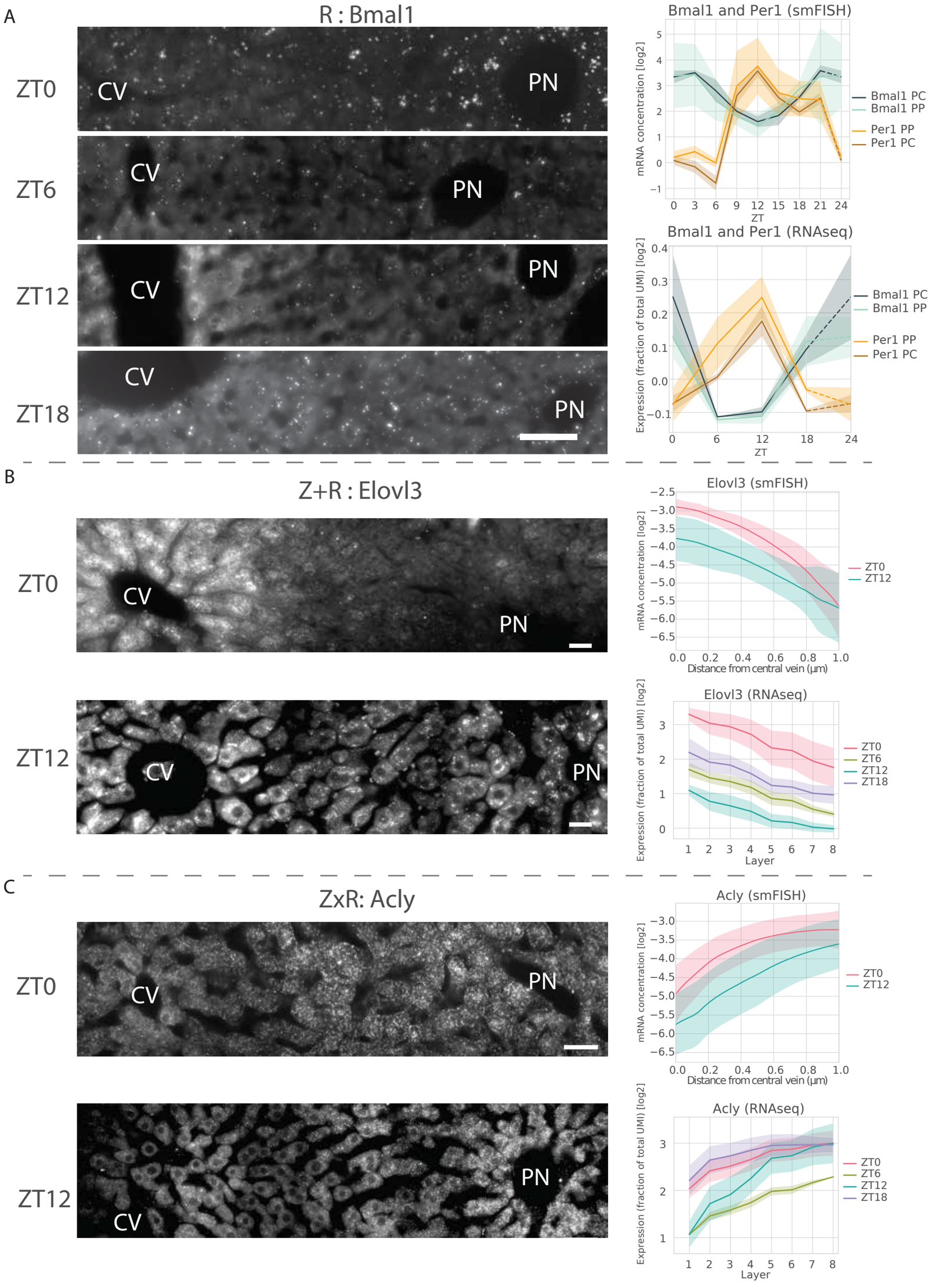

A
R : Bmal1
ZT0
ZT6
ZT12
ZT18
CV
PN
Bmal1 and Per1 (smFISH)
mRNA concentration [log2]
ZT
Bmal1 PC
Bmal1 PP
Per1 PP
Per1 PC
Bmal1 and Per1 (RNAseq)
Expression (fraction of total UMI) [log2]
B
Z+R : Elovl3
Elovl3 (smFISH)
Distance from central vein (µm)
ZT0
ZT12
Elovl3 (RNAseq)
ZT6
ZT18
Layer
C
ZxR: Acly
Acly (smFISH)
Acly (RNAseq)

Figure 5

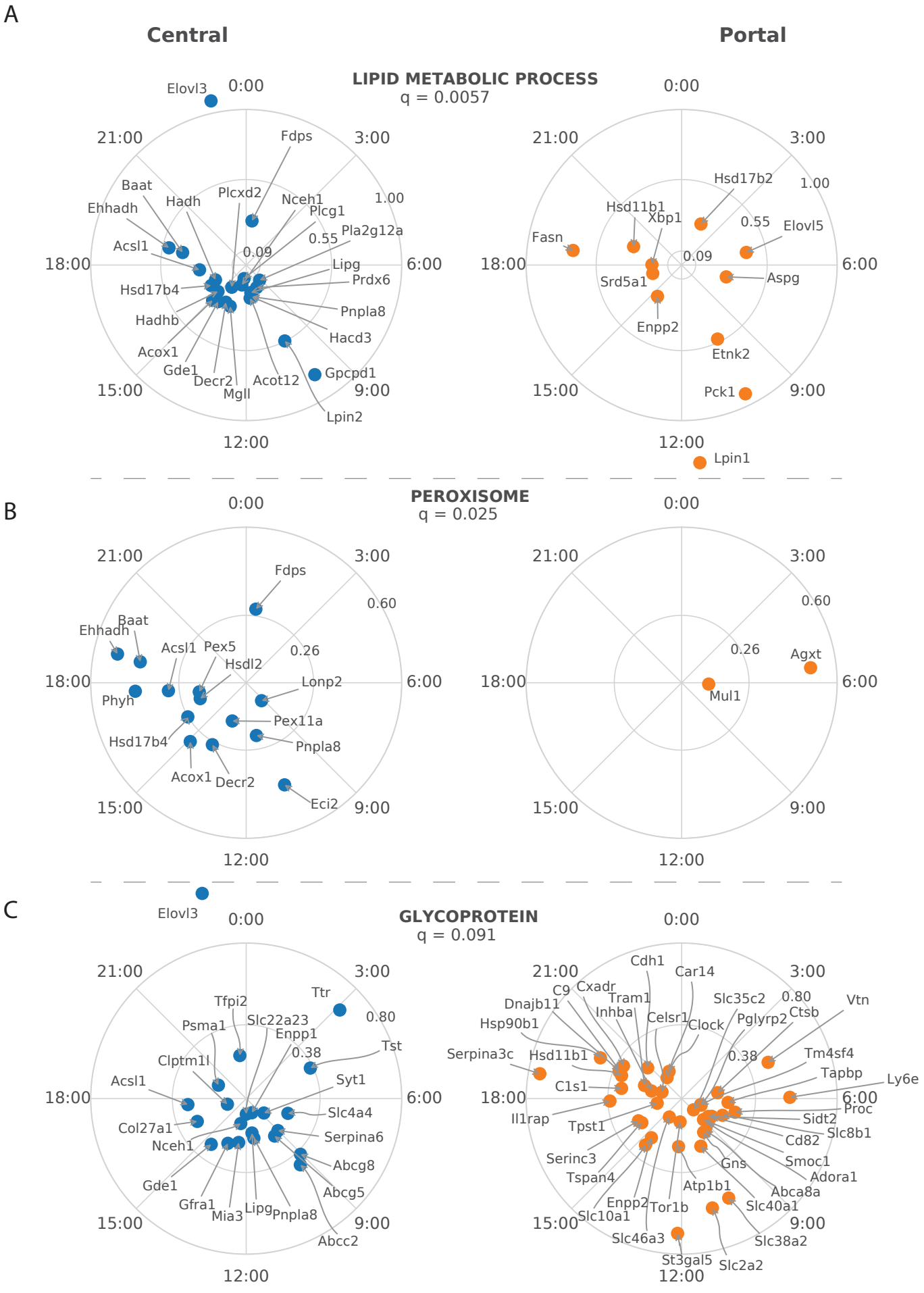
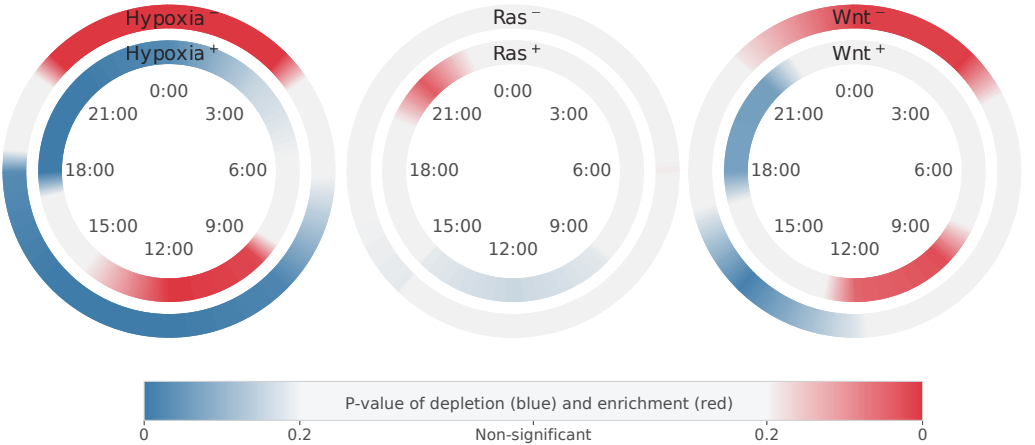
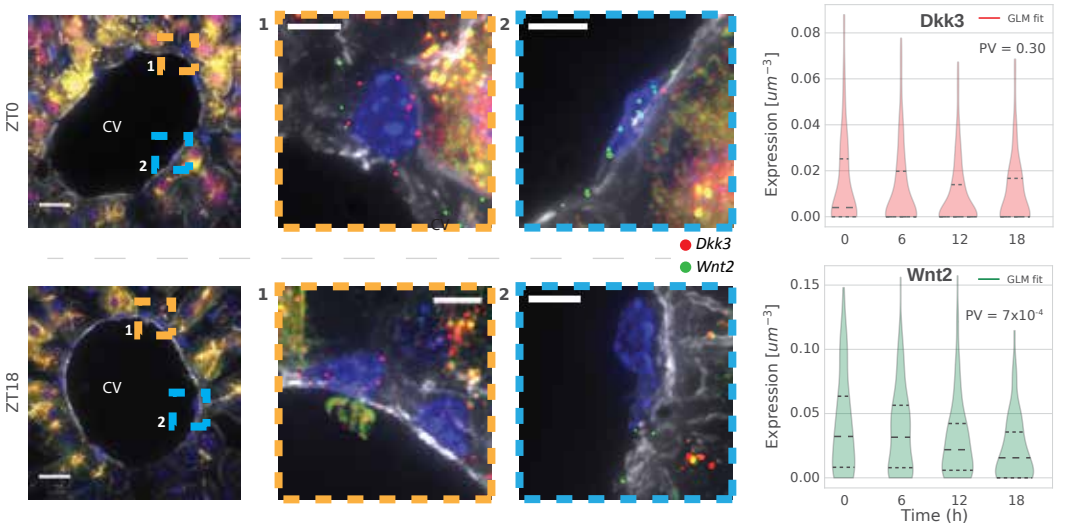

A
Central
Portal
LIPID METABOLIC PROCESS
q = 0.0057
0:00
3:00
6:00
9:00
12:00
15:00
18:00
21:00
1.00
0.55
0.09
Elovl3
Fdps
Plcxd2
Nceh1
Plcg1
Pla2g12a
Lipg
Prdx6
Pnpla8
Hacd3
Gpcpd1
Lpin2
Acot12
Mgll
Decr2
Gde1
Acox1
Hadhb
Hsd17b4
Acsl1
Ehhadh
Baat
Hadh
Hsd17b2
Elovl5
Aspg
Etnk2
Pck1
Lpin1
Enpp2
Srd5a1
Xbp1
Hsd11b1
Fasn
B
PEROXISOME
q = 0.025
0.60
0.26
Fdps
Baat
Ehhadh
Acsl1
Pex5
Hsdl2
Lonp2
Pex11a
Pnpla8
Eci2
Decr2
Acox1
Hsd17b4
Phyh
Agxt
Mul1
C
GLYCOPROTEIN
q = 0.091
0.80
0.38
Elovl3
Ttr
Tfpi2
Psma1
Slc22a23
Enpp1
Tst
Syt1
Slc4a4
Serpina6
Abcg8
Abcg5
Abcc2
Pnpla8
Lipg
Mia3
Gfra1
Gde1
Nceh1
Col27a1
Acsl1
Clptm1l
Cdh1
Car14
Cxadr
C9
Tram1
Dnajb11
Inhba
Celsr1
Clock
Slc35c2
Pglyrp2
Ctsb
Vtn
Hsp90b1
Serpina3c
Hsd11b1
Tm4sf4
Tapbp
Ly6e
C1s1
Il1rap
Tpst1
Proc
Sidt2
Slc8b1
Cd82
Smoc1
Adora1
Gns
Serinc3
Tspan4
Atp1b1
Abca8a
Slc40a1
Enpp2
Tor1b
Slc10a1
Slc46a3
Slc38a2
St3gal5
Slc2a2

Figure 6

A


Hypoxia −
Hypoxia +
Ras −
Ras +
Wnt −
Wnt +
0:00
3:00
6:00
9:00
12:00
15:00
18:00
21:00
P-value of depletion (blue) and enrichment (red)
0
0.2
Non-significant
0.2
0

B


ZT0
ZT18
CV
1
2
Dkk3
Wnt2
GLM fit
PV = 0.30
PV = 7x10⁻⁴
Expression [um⁻³]
Time (h)
0
6
12
18