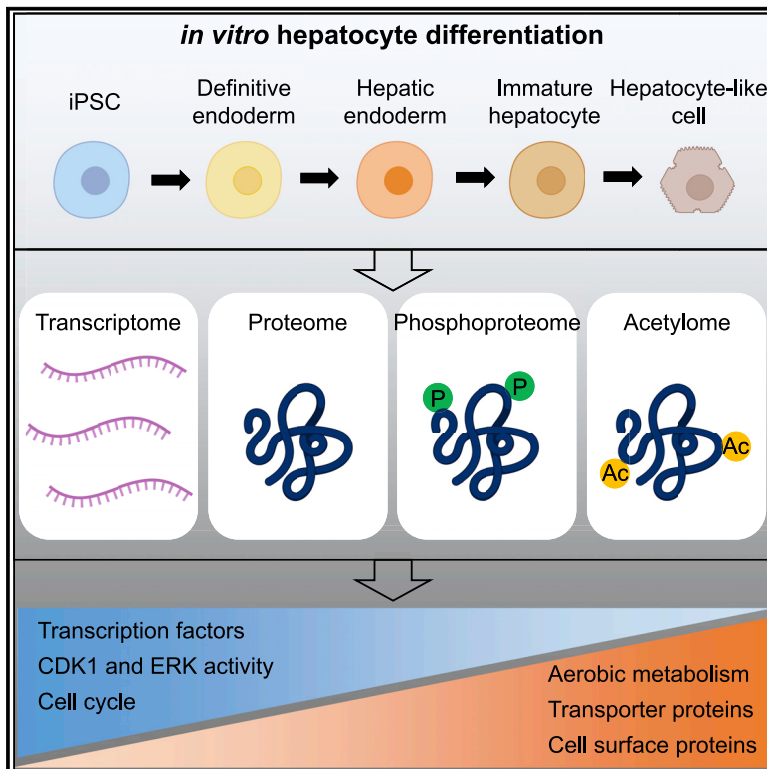# High temporal resolution proteome and phosphoproteome profiling of stem cell-derived hepatocyte development

Graphical abstract

Highlights

- High temporal resolution proteome of stem cell-derived hepatocyte differentiation

- *In vitro* hepatocyte differentiation is accompanied by a major molecular switch

- 3D liver organoids better resemble *in vivo* hepatocytes than 2D models

- Stem cell-derived hepatocytes are fetal-like and not fully ADME/tox compatible

Authors

Johannes Krumm, Keisuke Sekine, Patroklos Samaras, ..., Barbara Treutlein, J. Gray Camp, Bernhard Kuster

Correspondence

kuster@tum.de

In brief

Krumm et al. provide an extensive resource of the temporal changes in the proteome, phosphoproteome, and acetylproteome during induced pluripotent stem cell (iPSC)-derived hepatocyte differentiation. They discover a major molecular switch and find that *in vitro* differentiated hepatocytes are in a fetal-like state with decreased expression of important ADME-tox proteins.

CellPress

Cell Reports

Resource

High temporal resolution proteome and phosphoproteome profiling of stem cell-derived hepatocyte development

Johannes Krumm,[1] Keisuke Sekine,[2,3] Patroklos Samaras,[1] Agnieska Brazovskaja,[4] Markus Breunig,[5] Ryota Yasui,[3] Alexander Kleger,[5] Hideki Taniguchi,[3,6] Mathias Wilhelm,[1,7] Barbara Treutlein,[8] J. Gray Camp,[9] and Bernhard Kuster[1,10,11,*]

[1]Chair of Proteomics and Bioanalytics, Technical University of Munich, 85354 Freising, Germany
[2]Laboratory of Cancer Cell Systems, National Cancer Center Research Institute, Tokyo 104-0045, Japan
[3]Department of Regenerative Medicine, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa 236-004, Japan
[4]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany
[5]Department of Internal Medicine I, Ulm University Hospital, 89081 Ulm, Germany
[6]Division of Regenerative Medicine, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan
[7]Computational Mass Spectrometry, Technical University of Munich, 85354 Freising, Germany
[8]Department of Biosystems Science and Engineering, ETH Zurich, 4058 Basel, Switzerland
[9]Institute of Molecular and Clinical Ophthalmology Basel, 4056 Basel, Switzerland
[10]Bavarian Biomolecular Mass Spectrometry Center (BayBioMS), Technical University of Munich, 85354 Freising, Germany
[11]Lead contact
*Correspondence: kuster@tum.de
https://doi.org/10.1016/j.celrep.2022.110604

SUMMARY

Primary human hepatocytes are widely used to evaluate liver toxicity of drugs, but they are scarce and demanding to culture. Stem cell-derived hepatocytes are increasingly discussed as alternatives. To obtain a better appreciation of the molecular processes during the differentiation of induced pluripotent stem cells into hepatocytes, we employ a quantitative proteomic approach to follow the expression of 9,000 proteins, 12,000 phosphorylation sites, and 800 acetylation sites over time. The analysis reveals stage-specific markers, a major molecular switch between hepatic endoderm versus immature hepatocyte-like cells impacting, e.g., metabolism, the cell cycle, kinase activity, and the expression of drug transporters. Comparing the proteomes of two- (2D) and three-dimensional (3D)-derived hepatocytes with fetal and adult liver indicates a fetal-like status of the *in vitro* models and lower expression of important ADME/Tox proteins. The collective data enable constructing a molecular roadmap of hepatocyte development that serves as a valuable resource for future research.

INTRODUCTION

Hepatocytes are parenchymal cells of the liver and constitute about 70%–80% of its mass. They execute most liver functions, such as production of bile, secretion of albumin, metabolism of carbohydrates and fatty acids, and detoxification of xenobiotics. Because of the latter, hepatocytes serve as an important model system in drug discovery and development. Primary human hepatocytes (PHHs) have become the gold standard for the assessment of drug metabolism and toxicity. Although PHHs reduce the need for animal testing, their shortage, potential changes during *in vitro* culture, and high donor-to-donor variability limit their broader applicability. To overcome the lack of expandable and standardized sources of hepatocytes, differentiated hepatocyte-like cells derived from human induced pluripotent stem cells (iPSCs) are increasingly explored as alternatives (Davidson et al., 2015; Bartfeld and Clevers, 2017). iPSCs are reprog-

rammed somatic cells that possess the ability to propagate indefinitely and that are able to differentiate into cells of all three germ layers. For differentiation into hepatocytes, iPSCs are treated with Wnt3A and activin A to promote definitive endoderm (DE) differentiation, eventually leading to the formation of, e.g., the liver, pancreas, and intestines. Further differentiation into hepatocytes can be triggered by the addition of hepatocyte growth factor (HGF), dexamethasone (DEX), and oncostatin M (OSM). The ability to control this process has enabled studying human developmental biology and diseases (Karagiannis et al., 2019). Further advances in cell culture have led to the generation of liver buds (LBs), three-dimensional (3D) tissue-like structures composed of iPSC-derived hepatic endodermal (HE), mesenchymal, and endothelial cells (ECs). On transplantation into mice, LBs form vascularized and functional tissue that, e.g., prolonged the lives of mice in a model of liver failure, indicating the potential value of LBs for regenerative medicine (Takebe et al., 2013).

The molecular processes underlying hepatocyte differentiation from iPSCs are not well understood. Several studies have described the differentiation from iPSCs to hepatocytes, but their characterization was mostly limited to a small number of marker proteins (Vyas et al., 2018; Takebe et al., 2017; Sekine et al., 2020) or mRNA expression analysis (Ang et al., 2018). Camp et al. (2017) used single-cell transcriptomics (single-cell RNA sequencing [scRNA-seq]) to investigate cell lineage progression of pluripotent cells toward hepatocytes and benchmarked iPSC-derived liver organoids to fetal cells, finding a striking similarity. Because most biological processes are regulated at the protein level and post-translational modifications (PTMs) are often required to control metabolic and other cellular activities, analyzing iPSC differentiation at the proteome level is a logical next step. Few such experiments have been performed so far, and insights have either been restricted to a more general and early embryonic developmental stage (Van Hoof et al., 2009; Yang et al., 2019; Rigbolt et al., 2011), limited because of low proteome coverage (Baxter et al., 2015; Rowe et al., 2013), or non-exhaustive analysis of the proteomic data (Hurrell et al., 2019). Here, we systematically measured very deep and quantitative temporal profiles of proteome expression changes, phosphorylation dynamics, and signatures of protein acetylation from iPSCs to hepatocytes in a two-dimensional (2D) cell culture model and integrated published scRNA-seq data collected for the same system (Camp et al., 2017). This revealed, for instance, stage-specific markers and changes in signaling pathways and allowed the identification of candidate substrates of kinases and histone deacetylases. A comparison of hepatocytes differentiated in 2D or 3D with PHHs and fetal liver delineated clear molecular differences, thereby laying the foundation to develop strategies for an improved functional correspondence of *in vivo* liver cells and iPSC-derived hepatocytes.

RESULTS

Time-resolved proteome, phosphoproteome, and acetylome profiles of iPSC differentiation into hepatocytes

We followed a published protocol that uses TkDA3–4 iPSCs isolated from human skin and performed two independent hepatocyte differentiations (Camp et al., 2017). Samples were collected for proteomic analysis (expression, phosphorylation, acetylation) on day (d) 0 (iPSC), d6 (DE), d8 (HE), d13 (immature hepatocyte-like [IH]), and d21 (mature hepatocyte-like [MH]; Figures 1A and S1). Samples representing each stage were stable isotope labeled using tandem mass tags (TMTs) and analyzed by quantitative mass spectrometry (MS; Figure 1B). The temporal profiles covered ~9,000 proteins, ~12,000 phosphorylation sites (P-sites), and ~800 acetylation sites (Ac-sites), providing a comprehensive proteomic map of hepatocyte differentiation (Figure 1C). Principal-component analysis (PCA) of quantified proteins revealed a time-dependent separation, confirming that the proteomic data recapitulated the cellular transition along the differentiation process (Figure 1D). As Camp et al. (2017) demonstrated previously using scRNA-seq data on the same system, nearly 100% of cells reached the DE stage and 90% of cells expressed hepatoblast-like marker proteins at the IH

stage. Therefore, bulk measurements at the proteome level appear sufficient to recapitulate protein expression for the majority of cells at each specific stage. The high quality of the proteomic data was confirmed by comparing protein expression of accepted markers with quantitative PCR data collected from the same samples and (averaged) scRNA-seq data from the aforementioned study (Figure 1E) (Camp et al., 2017). The three data layers show consistent results for a number of stage-specific markers. However, the global correlation between mRNA and protein levels was rather weak at all five stages (Figure S2A) (Liu et al., 2016; Wang et al., 2019). We next calculated the fold-change (FC)/ratios of proteins and transcripts between adjacent stages of differentiation to ask the question if similar conclusions could be drawn from the proteome or RNA-seq data. Interestingly, these data correlated even worse, suggesting that the temporal dynamics are very discrepant between transcripts and protein levels (Figure S2B). It is noteworthy that the correlation of protein and mRNA between the stages that are furthest apart in time (iPSC and MH) is better (Figure S2C; Pearson correlation of 0.59). This indicates that mRNA levels can be used as proxies for protein levels to some extent but are insufficient to reveal temporal dynamics at the protein level. This interpretation is illustrated by plotting the temporal profiles of all differentially expressed transcription factors (TFs). Although similar overall trends from iPSC to MH can be observed at both mRNA and protein levels, the proteomic data bring out the temporal aspect more clearly (Figure S2D).

The transition of HE cells to IH cells is accompanied by a distinct metabolic switch

Analysis of variance (ANOVA) revealed a large number of proteins with statistically significant expression changes across the differentiation process (false discovery rate [FDR] < 0.05; FC > 2). These 4,956 proteins were hierarchically clustered into 10 distinct expression profiles (Figure 2A), of which clusters 3 and 10 highlight the major protein expression changes occurring between HE and IH. Overrepresentation analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways revealed enrichment of multiple metabolic pathways in cluster 3 (Figure S2E). Notably, all 13 detected proteins of the tricarboxylic acid (TCA) cycle showed a synchronous expression pattern demonstrating their concerted upregulation (Figure 2B). Again, this temporal behavior was more robustly observed at the protein level (Figure S2F). In particular, the observed upregulation of succinate dehydrogenase (SDHA/B), which connects the TCA cycle with the respiratory electron transfer chain, together with an increased expression of multiple respiratory enzymes, indicates that cellular metabolism switches to oxidative phosphorylation in IH cells (Figure S2G). This is supported by observed elevated levels of proteins associated with the mitochondrial ribosome (Figure 2C, left panel), which translates proteins of the electron transport chain (Figure S2G). Cytoplasmic ribosomal proteins showed the opposite effect (Figure 2C, right panel). All three subunits of the pyruvate dehydrogenase complex were upregulated from HE to IH, suggesting an increased conversion of pyruvate into acetyl-CoA (Figure S2H), which can subsequently be metabolized in the TCA cycle. Another highly enriched pathway was peroxisome metabolism, which plays a key role in fatty acid
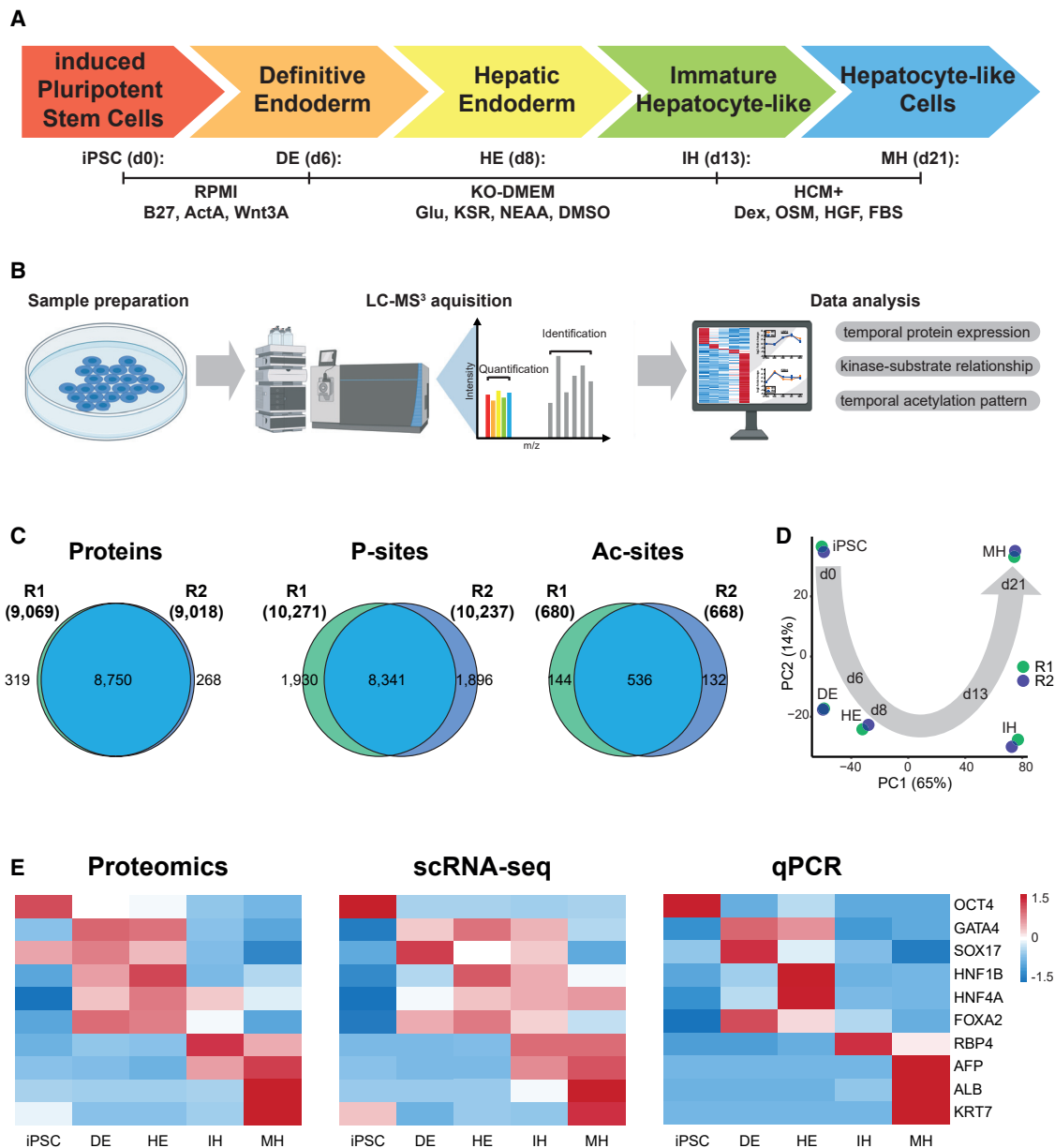
A

B

C

D

E

Figure 1. Temporal proteome, phospho-proteome, and acetyl-proteome profiling of iPSC-derived hepatocyte-like cells

(A) Schematic overview of the procedure for differentiating human iPSCs into hepatocyte-like cells. Day (d) 0 to d21 denote the time in days at which samples were collected. Information provided below the time axis refers to cell culture media and supplements.

(B) Schematic overview of the proteomic workflow and data analysis. Samples collected at each time point of cell culture were encoded by stable isotopes (TMT), combined, and analyzed by MS to generate temporal profiles for further data analysis.

(C) Venn diagrams showing the number of identified proteins, P-sites, and Ac-sites from both independent replicates (R1, R2).

(D) PCA of the proteomic data.

(E) Z-scored normalized heatmaps of the quantities of known stage-specific differentiation markers. Left panel: proteomics data; middle panel: previously published scRNA-seq data (Camp et al., 2017); right panel: qPCR data.

See also Figure S1.

metabolism (Figures S2E and S2I, lower panel). Specifically, the observed increased mRNA expression of the TF *PPARA* (Figure S2I, upper panel) led to higher protein expression of its target genes (Figure 2D), pointing to an important role for PPARA in the metabolic switch from HE to IH.

Phosphorylation changes regulating the cell cycle precede dynamics on the proteome level

KEGG analysis also revealed an enrichment of cell-cycle-related proteins in cluster 10 (Figure S2E). Although mRNA levels of genes involved in DNA replication decreased between iPSC
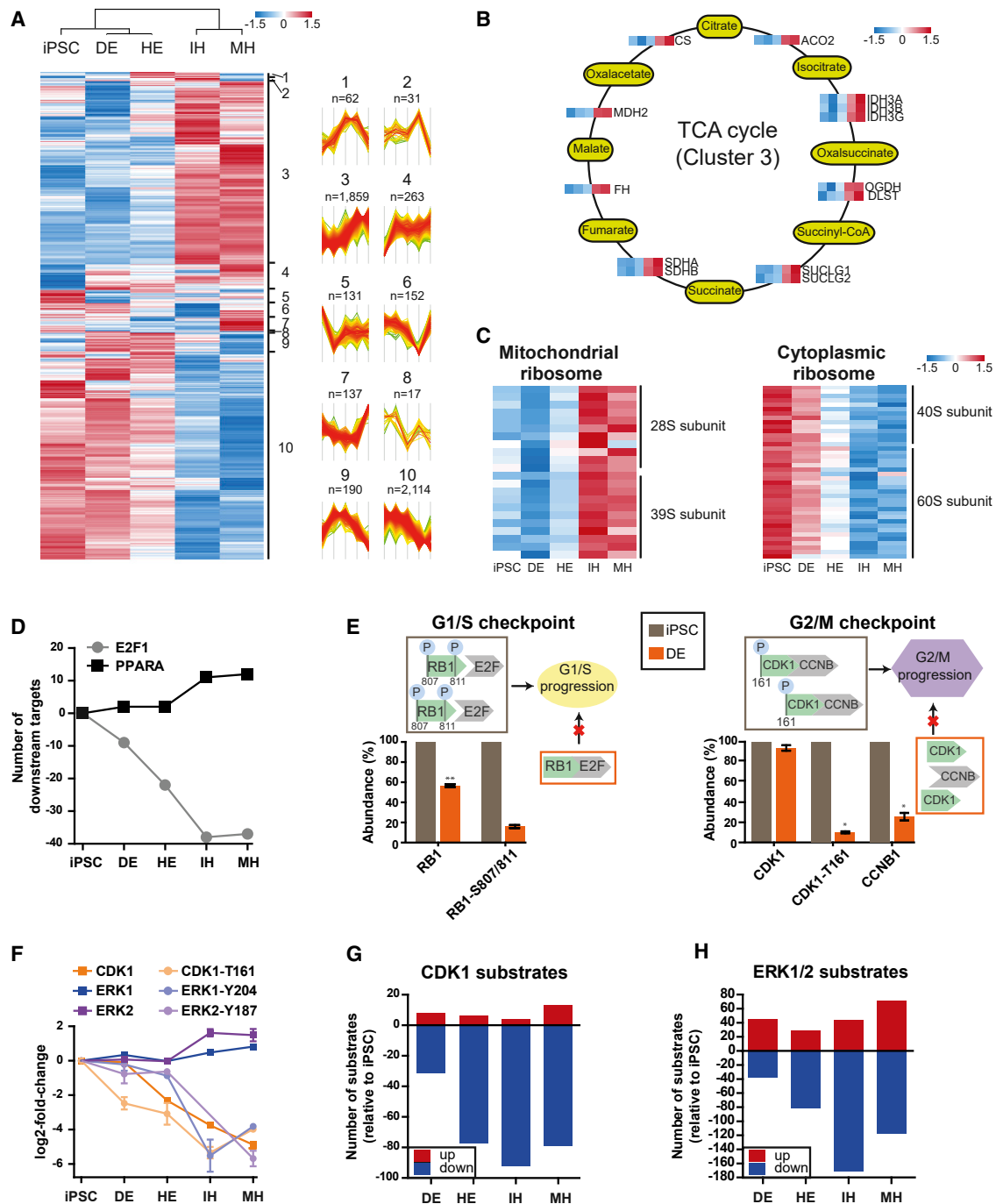
Figure 2. Exploring protein and phosphoprotein dynamics during hepatocyte differentiation

(A) Heatmap showing results of hierarchical clustering of 4,956 differentially expressed proteins (ANOVA, BH corrected: FDR < 0.05 and FC > 2 at minimum one time point). The right panel shows the expression profiles of proteins in all 10 clusters (n indicates the number of proteins in the cluster).

(B) Expression of proteins (Z scored) related to the tricarboxylic acid (TCA) cycle.

(C) Expression of proteins (Z scored) associated with the mitochondrial (left) or cytoplasmic (right) ribosome.

(D) Number of significantly (ANOVA, BH corrected: FDR < 0.05 and FC > 2) temporally regulated targets of PPARA and E2F1.

(E) Left panel: model showing how protein expression and phosphorylation levels of the G1/S checkpoint regulator RB1 control cell-cycle progression. Phosphorylation of S807/S811 of RB1 impairs binding to E2F TFs. This enables E2F to translocate into the nucleus and drive G1/S progression. In contrast, unphosphorylated RB1 binds E2F and inhibits G1/S progression. Right panel: G2/M checkpoint regulation by CDK1 and CCNB1. Phosphorylation of T161 of CDK1 leads to CCNB1 binding and activation of the kinase, leading to G2/M progression. The y axes show protein abundance of DE relative to iPSCs. Columns represent the average of two replicates, and error bars show the range. Asterisks denote significance (ANOVA, BH corrected: *FDR < 0.05; **FDR < 0.01).

(legend continued on next page)

and DE, the decline in the respective protein levels was observed only later (from HE to IH; Figure S2J). In contrast, we observed a reduction of phosphorylation of RB1-S807/811 and CDK1-T161 from iPSC to DE, suggesting an early post-translational regulation of the cell cycle by G1/S and G2/M checkpoint inhibition. Hyperphosphorylated RB1 is unable to bind the G1/S progressing TF family E2F, whereas the unphosphorylated protein binds E2F, thereby impairing cell-cycle progression. To confirm this hypothesis, we analyzed the expression of significantly changing proteins that are regulated by E2F1, and we observed a steady decrease along the differentiation timeline pointing to progressively decreasing E2F1 activity caused by RB1 binding (Figures 2D and 2E, left panel).

The G2/M checkpoint is the second key regulator of the cell cycle and mainly controlled by CDK1. Although its protein expression did not change between iPSC and DE, we observed a significant reduction of T161, indicating lower CDK1 activity (Figures 2E and 2F) (Qu et al., 2021; Coulonval et al., 2011). In addition, levels of CCNB1, the major cyclin for CDK1, were strongly decreased between iPSC and DE (Figure 2E, right panel). Further support for lower CDK1 activity comes from the observation that the phosphorylation levels of a large number of known and predicted CDK1 substrates decreased during differentiation (Figure 2G, blue bars; Table S1). We note that CDK1 protein levels also decreased from DE to very low levels at MH (Figure 2F), consistent with reduced proliferation accompanying cellular maturation (Estefania et al., 2012; Boward et al., 2016). Albeit to a lesser extent, similar observations were made for ERK1 and ERK2. Although protein levels increased somewhat during differentiation, the activity of the two kinases was substantially reduced (indicated by the 4- to 16-fold loss of Y187 and Y204 intensity; Figure 2F). A reduction in phosphorylation levels was also observed for multiple known or predicted ERK1/2 substrates (Figure 2H, blue bars; Table S1). It should be noted that the ERK1/2 phosphorylation motif is not as selective as the CDK1 motif, potentially accounting for the more diffuse ERK1/2 substrate analysis observed during *in vitro* differentiation. Many more candidate kinase-substrate relationships may be derived from the data and highlight the value of the work as a molecular resource.

Temporal proteomic and phosphoproteomic profiles are consistent between iPSC donors

To test the validity of the above results, we repeated the analysis using a different iPSC donor (Ff-I01) and a slightly modified protocol (Sekine et al., 2020). Here, three independent differentiations were performed, and the temporal resolution was increased to 12 time points (Figure S3A). Using the same proteomic workflow as above, we identified ∼9,000 proteins and ∼12,000 P-sites (Figure S3B). Samples from d0 (iPSC), d6 (DE), d10 (HE), d13 (IH), and d21 (MH) were employed to test for consistency with the first differentiation. Proteome and transcriptome analysis confirmed successful differentiation, the PCA analysis showed the expected time-dependent separation (Figures S3C and S3D), and hierarchical clustering of differentially expressed proteins (n = 5,105) showed similar dynamics for metabolic pathways and cell-cycle-related proteins as before (Figures S3E–S3G). The analysis also confirmed the above effects on protein and phosphorylation levels of phosphorylated RB1, as well as reduction of phosphorylation of CDK1 and ERK1/2 substrates (Figures S3H and S3I; Table S1).

Differentiation of iPSCs into hepatocytes is associated with opposing but coordinated expression changes in functional protein classes

Aiming for a more functional analysis, we divided the temporal proteome profiles into groups of proteins with distinct biological functions. Although the expression of epigenetic modifiers (defined according to Medvedeva et al., 2015) decreased, the levels of cell surface proteins increased during hepatic differentiation of both iPSC lines (Figure 3A). Among proteins derived from the cell surface protein atlas (Bausch-Fluck et al., 2015), lysosomal proteins were overrepresented in both biological replicates (Figures 3B and S4A). Beyond their catabolic functions, lysosomes are increasingly recognized as important organelles for differentiation processes (Monteleon et al., 2018; Kobayashi et al., 2019; Liang et al., 2020; Villegas et al., 2019; Ferguson, 2015), and many of the changes observed in the proteome may reflect such non-canonical functions. Of the 726 detected cell surface proteins, 97 were differentially expressed at one or more time points along the differentiation in both iPSC lines (Table S2). Interestingly, the two fibroblast growth factor receptors, FGFR1 and FGFR4, were significantly upregulated in HE cells, indicating elevated FGF signaling, which was shown to be an important pathway for early hepatocyte development (Jung et al., 1999; Berg et al., 2007). This list provides marker proteins that may be used for controlling future differentiations and for purifying specific cell populations.

Most proteins annotated as epigenetic modifiers showed a strong downregulation along the differentiation timeline (88% in TkDA3–4 cells and 71% in Ff-I01 cells; Figure 3A), suggesting a global epigenetic rewiring of the cells. That said, diametrically opposing trends were observed for the protein deacetylases SIRT1 and SIRT2, in line with studies that correlate high levels of SIRT1 with pluripotency (Calvanese et al., 2010; Lee et al., 2012) and increasing levels of SIRT2 with differentiation (Si et al., 2013; Cha et al., 2017) (Figures 3C and S4B). Given that SIRT2 was the only significantly upregulated deacetylase between HE and IH, we speculated that any presumed increase in SIRT2 activity may be measured by decreasing levels of protein acetylation. To address this, we enriched acetylated peptides via immunoprecipitation (IP) and indeed detected differentially regulated acetylated lysine residues between HE and IH. Among these was PGAM-K100, a known substrate of

---

(F) Temporal protein expression of ERK1/2 and CDK1 and levels of P-sites indicative of kinase activity. Data are normalized to iPSCs and represent the average log$_2$-FC of two replicates, and error bars depict the range.

(G) Number of significantly upregulated/downregulated CDK1 substrates relative to iPSCs.

(H) Same as in (G) but for ERK1/2 substrates.
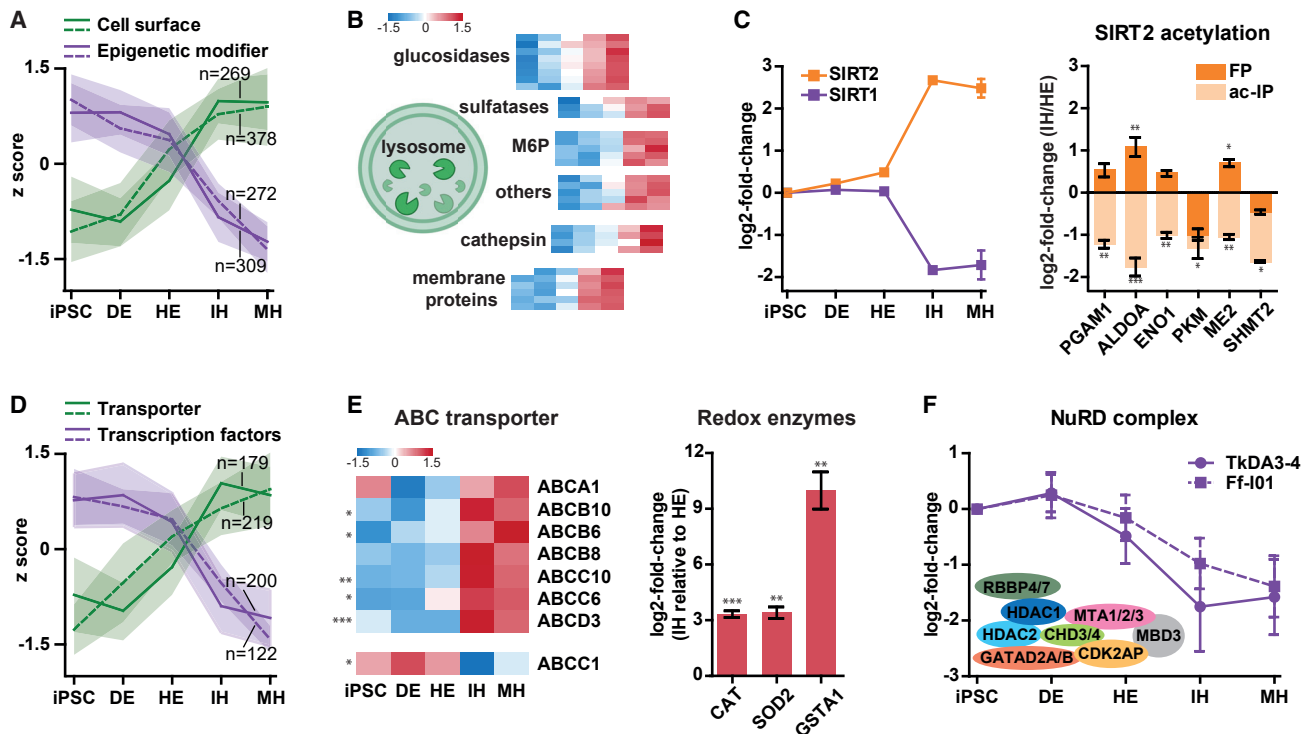
See also Figure S2 and Table S1.

Figure 3. Temporal dynamics of selected functional protein classes

(A) Temporal expression profiles of cell surface proteins and epigenetic modifiers. Proteins were hierarchically clustered (n = 8 clusters), and the median and standard deviation of the top cluster are shown as lines and range, respectively. Solid lines represent results for TkDA3–4 cells and dashed lines results for Ff-I01 cells.

(B) Z-scored expression of lysosomal proteins (TkDA3–4 cells), a subset of the "cell surface" category in (A).

(C) Left panel: expression of SIRT1/2 relative to iPSCs. Right panel: change of protein expression (FP) and acetylation level changes on the same proteins (ac-IP, light orange) of known and putative SIRT2 substrates. Error bars depict the range, and asterisks denote significance (ANOVA, BH corrected: *FDR < 0.05; **FDR < 0.01; ***FDR < 0.001).

(D) Same as in (A) but for transporter proteins and TFs.

(E) Left panel: Z-scored heatmap of normalized ABC transporter expression in TkDA3–4 cells. Right panel: expression differences of reactive oxygen species-reducing enzymes between HE and IH. Error bars depict the range, and asterisks denote significance (ANOVA, BH corrected: *FDR < 0.05; **FDR < 0.01; ***FDR < 0.001).

(F) Dynamic expression of the NuRD complex. Data points denoting the median expression of identified NuRD complex members relative to iPSCs. Error bars represent the standard deviation.

See also Figure S4.

SIRT2 with important roles in energy production (Xu et al., 2014). Similarly, acetylation of the functionally related enzymes and known SIRT2 substrates ALDOA, ENO1, and PKM (Cha et al., 2017) were downregulated. Our analysis revealed several acetylated lysine residues on these proteins together with their dynamics during hepatocyte differentiation that, together with ME2 and SHMT2, likely represent SIRT2 substrates (Figure 3C, right panel).

The liver is a molecular powerhouse, synthesizing and secreting large quantities of biomolecules, such as bile, proteins, and lipids. As a result, hepatocytes express high levels of transporter proteins (Figure 3D). One important group is ATP-binding cassette (ABC) transporters, which are also tightly linked to drug metabolism, efflux, and resistance (Vasiliou et al., 2009). Although the levels of multiple members of this family increased over time, ABCC1, a protein that positively regulates glutathione efflux (Wiese and Stefan, 2019), was strongly downregulated

(Figures 3E and S4C, left panels). Decreased glutathione efflux together with the observed higher expression of enzymes involved in eliminating reactive oxygen species illustrates the cellular adaptation to the increased use of oxidative phosphorylation for energy production in hepatocytes that would be expected to go along with higher oxidative stress (Figures 3E and S4C, right panels).

We found that 76% (for TkDA3–4) and 51% (for Ff-I01) of all detected TFs were expressed at high levels in the pluripotent state but at far lower levels on maturation (Figure 3D). This likely reflects the need to upkeep a large number of transcriptional programs in the pluripotent state. Once these are no longer needed, the respective TFs are degraded. One example is the nucleosome remodeling deacetylase (NuRD) complex (Figure 3F). It forms a transcriptional regulator with important roles in lineage commitment such that it maintains transcriptional heterogeneity in pluripotent stem cells enabling subsequent development into
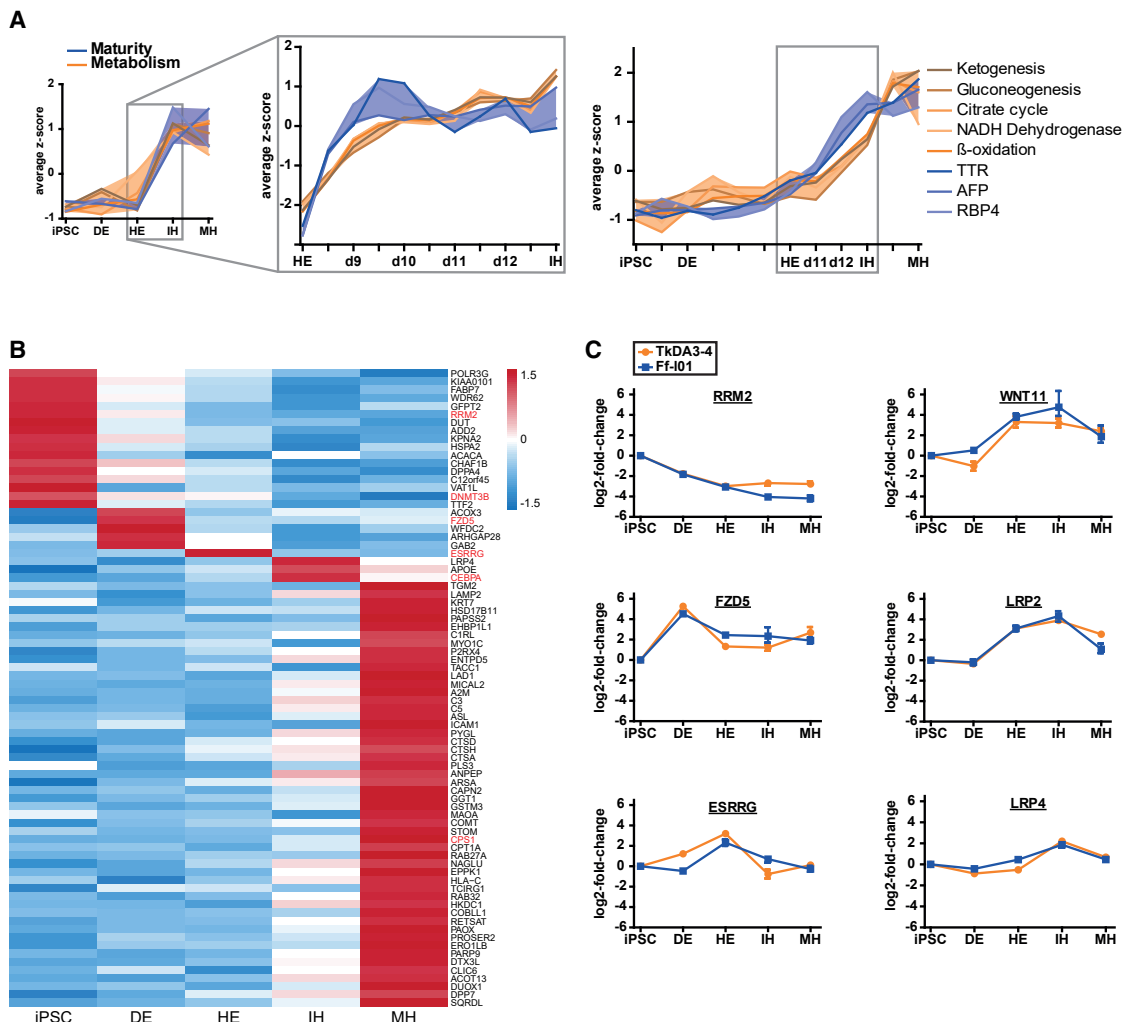
Figure 4. Identification of stage-specific marker proteins during hepatocyte differentiation

(A) Temporal expression profile of selected maturation marker (blue) and metabolism pathways (orange) from the TkDA3–4 cells (left panel), a corresponding higher-resolution experiment (middle panel), and the differentiation with Ff-I01 cells (right panel). Line represents the median *Z* score of all proteins related to each pathway.

(B) *Z*-scored heatmap representation of stage-specific protein marker.

(C) Dynamic expression of Wnt-related proteins relative to iPSCs. Data points represent the average and error bars depict the range.

See also Figure S5 and Table S2.

multiple different cell types (Reynolds et al., 2012; Basta and Rauchman, 2015). Interestingly, the SIN3-complex, which, like the NuRD complex, belongs to the family of HDAC1/2 complexes, showed a similar dynamic (Figure S4D).

Hepatocyte maturation precedes the metabolic switch

With the apparent upregulation of multiple metabolic pathways, we sought to shed more light on the temporal correlation of hepatocyte maturation and the metabolic switch by increasing temporal resolution. Hence we repeated the experiment taking samples every 12 h between HE and IH (Figure 4A). The results indicate that the upregulation of maturation markers such as TTR, AFP, and RBP4 precede the changes in metabolism in TkDA3–4 cells (Figure 4A, middle panel). This interpretation is

supported by similar results from experiments using Ff-I01 cells, suggesting that the metabolic switch is a result of maturation and not vice versa (Figure 4A, right panel).

Temporal proteome profiling reveals stage-specific markers

Given the depth of the proteomic data, we sought to identify stage-specific marker proteins by filtering differentially expressed proteins (ANOVA, FDR < 0.05) for at least 2-fold higher abundance at one time point compared with all other time points. This led to a list of 78 stage-specific proteins that exhibited similar expression profiles in both iPSC lines (Figure 4B; Table S2). These proteins may provide a starting point for further investigations into the differentiation process or serve as a reference list for developing
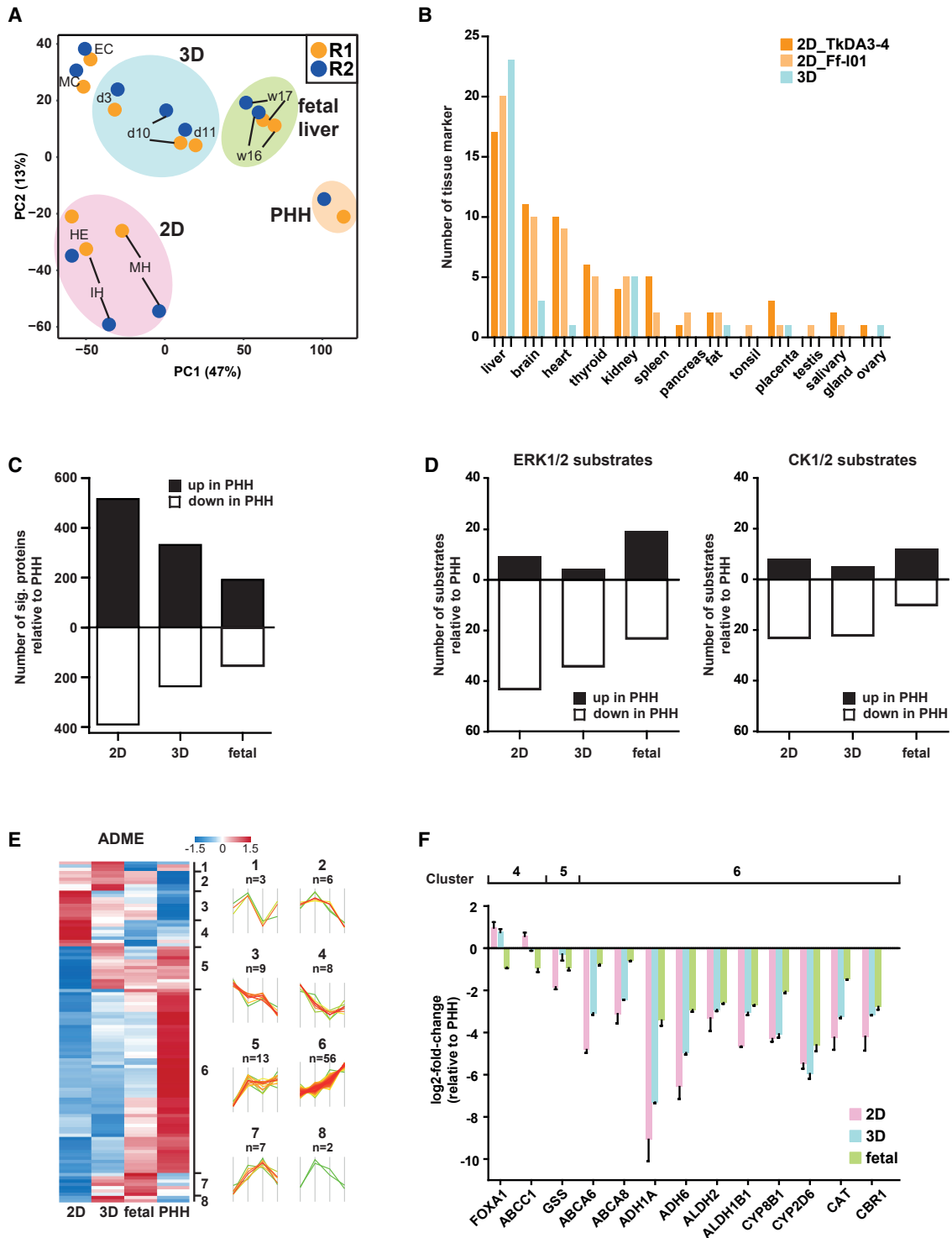
Figure 5. Comparison of *in vitro* and *in vivo* hepatocytes

(A) PCA of the *in vitro* (2D, 3D, EC, MC) and *in vivo* (fetal liver, PHH) samples.

(B) Number of significantly (ANOVA, BH corrected: FDR < 0.05) upregulated tissue-specific protein markers (Wang et al., 2019) between HE and MH (for 2D) or d3 and d11 (for 3D), respectively.

(C) Number of significantly regulated (ANOVA, BH corrected: FDR < 0.05 and FC > 2) proteins relative to PHH.

(D) Same as (C), but showing the number of significantly regulated substrates of ERK1/2 (left panel) or CK1/2 alpha (right panel), respectively.

(legend continued on next page)

and evaluating future differentiation protocols. Most markers characterized iPSCs or MHs, whereas few were found for the intermediate stages. This may be related to cell-to-cell heterogeneity, which cannot be assessed by analyzing bulk proteomes. One example for an iPSC marker is ACACA, the rate-limiting enzyme in long fatty acid biosynthesis. High expression of this protein and fatty acid synthase (FASN) suggests elevated fatty acid synthesis in stem cells (Figures S5A and S5B). Another example is the DNA methyltransferase DNMT3B, which drastically decreased at the DE stage. The same was observed for DNMT1, which is required for methylation maintenance. DNMT3A did not show this behavior even though the two DNMT3 isoforms are thought to share common functions (Figure S5C) (Challen et al., 2014). In contrast with a study demonstrating that DNMT1, but not DNMT3A/B, is essential for cardiomyocyte and hematopoietic differentiation (Jackson et al., 2004), our observation of prolonged expression of DNMT3A past the iPSC stage may indicate a particular role for hepatocyte differentiation. However, such a hypothesis needs to be validated by additional experiments.

Prior work (Li et al., 2017) showed that stem cells and mature hepatocytes express E-cadherin. But rather than sustaining expression of this epithelial marker, iPSCs first undergo epithelial-mesenchymal transition (EMT) for DE development. Our proteomics and transcriptomics data revealed reduced levels of E-cadherin and increased levels of the mesenchymal marker N-cadherin at the DE stage, thus supporting the earlier study (Figure S5B, right panel). GAB2 is also associated with EMT in cancer (Ding et al., 2016) and was identified as a DE marker (Figure S5A). ESRRG was the only protein observed to be specifically and highly expressed at the HE stage. This nuclear receptor has no known ligand but is thought to be constitutively active. The protein is highly expressed in placenta, but not in most other tissues, pointing to specific functions in a developing organism. As described above, most TFs decreased on maturation, which was particularly pronounced between HE and IH (Figure 3D). Interestingly, the TF CEBPA was one of only three IH markers, but its close homolog CEBPB was not significantly regulated (Figures S5A and S5B). Both TFs have known roles in liver development (Westmacott et al., 2006; Shen et al., 2000), but different expression profiles suggest that they are not functionally redundant during hepatocyte differentiation.

Even though many hepatocyte-like proteins were already upregulated at the IH stage, 52 of the 78 markers uncovered here are MH specific. Furthermore, 19 are functionally related to the immune system, which is an important but often underestimated function of the liver (Zhou et al., 2016; Alper et al., 1969). One protein that stands out is CPS1, the enzyme of the first committing step in the urea cycle (16-fold higher expression at MH compared with other states; Figure S5A). The liver is the primary location for detoxifying urea, which may explain the strong upregulation of the protein at MH. Beyond this, the same protein, when secreted from the liver, has been shown to exert a cyto-

kine-like and immune-system-mediating liver-protective function in response to liver injury that is independent of its function in the urea cycle (Park et al., 2019b).

Wnt signaling plays a role in most steps of iPSC to hepatocyte differentiation

Canonical Wnt signaling is an essential driver of DE development (Toivonen et al., 2013; McLean et al., 2007), and the supplementation of iPSCs with Wnt3A or the Wnt activator CHIR99021 during the first 6 days drives DE formation (Figure 1A). The proteomic analysis revealed expression changes of multiple regulators of Wnt signaling, including marker proteins across all stages, indicating a tightly orchestrated interplay of Wnt signaling activators and inhibitors (Figure 4C). First, iPSCs showed high levels of the Wnt inhibitor RRM2, which dropped substantially at subsequent stages. In contrast, the Wnt receptor FZD5 showed an 8-fold increase in DE cells, suggesting FZD5 as a stage-specific marker for DE. At the same time, a strong downregulation of WNT5A (a ligand for the non-canonical pathway) (Sato et al., 2010) occurred, which suggests that the non-canonical Wnt signaling is repressed (Figure S5D, left panel). Beyond the DE stage, FZD5 levels dropped again accompanied with upregulation of the Wnt inhibitors ESRRG and WNT11 at the HE stage (Bisson et al., 2015; Kang et al., 2018), as well as LRP4 at the IH stage (Ahn et al., 2017), indicating that Wnt signaling is not sustained during further hepatocyte maturation steps (Toivonen et al., 2013; Lade and Monga, 2011). Multiple LRPs are known co-receptors for Wnt signaling, and thus the highly synchronous expression profile of WNT11 and LRP2 (Figures 4C and S5D, right panel) indicates a potential interplay, which needs further validation experiments. The Wnt pathway aside, we also identified multiple significantly altered levels of proteins and P-sites associated with known and essential signaling pathways for hepatocyte differentiation (Figure S5E; Table S3).

Hepatocytes differentiated in 3D more closely resemble liver cells than those obtained by 2D culture

The above results were obtained by iPSCs differentiated in 2D monolayer culture. An alternative is 3D LBs, organ-like structures consisting of cells from the HE stage, ECs, and mesenchymal cells (MCs), all of which are derived from iPSCs that self-organize into LBs on pooling (Takebe et al., 2014). To explore the similarities and differences between hepatocytes derived by 2D and 3D culture (both *in vitro*), as well as fetal liver and PHHs (both *in vivo*), we measured their (phospho-)proteomes side by side (two replicates; Figure S6A). This comparison included ~8,800 proteins and ~12,700 P-sites (Figure S6B). PCA analysis clearly separated 2D, 3D, fetal liver, and PHH, with the differentiation stages of the cell culture models indicating a maturation process toward the *in vivo* hepatocytes (Figure 5A). In accordance with data from the 2D analysis, we found multiple metabolic pathways elevated during LB maturation and proteins related to DNA replication downregulated (Figure S6C).

(E) Heatmap showing hierarchical clustering of all 104 significantly regulated (ANOVA, BH corrected: FDR < 0.05) proteins associated with ADME (Schroder et al., 2013). Right panel represents all eight clusters with their corresponding number of members.

(F) Bar graph of the relative protein expression to PHH. Bars denote mean values, and the error bars represent the range.

See also Figure S6 and Table S4.

We also gauged the differentiation specificity of the 3D LBs by comparing significantly upregulated proteins between d11 versus d3 of the LBs and MH versus HE of the 2D differentiation datasets, and compared these proteins with a proteomic analysis of 29 human tissues that defined a list of tissue-specific protein markers based on their abundance profiles (Wang et al., 2019). This analysis confirmed that the stage-specific marker proteins of 2D and 3D hepatocytes are closer to the human liver than any other organ (Figure 5B). More importantly, 3D marker proteins were more specific for the liver than 2D markers that were also found in several other tissues. We next compared *in vitro* and *in vivo* samples by counting all differentially expressed proteins relative to PHH. The clear trend toward diminishing differences from 2D to 3D to fetal liver suggested that 3D hepatocytes are molecularly closer to fetal or primary adult hepatic cells (Figure 5C). This trend was also obvious from comparing the abundances of the top 25 to 1,000 most abundant PHH proteins with those in the 2D to 3D and fetal liver datasets (Figure S6D). Still, it is apparent that 3D LBs do not completely recapitulate fetal, let alone adult, hepatocytes.

In addition to the different levels of protein expression, we found altered activities of multiple kinases between *in vitro* and *in vivo* samples. As described above, the number of phosphorylated ERK1/2 and CDK1 substrates decreased along the differentiation process (Figures 2G and 2H). However, compared with PHH, kinase activities were still elevated in both *in vitro* models (Figures S6E and 5D, left panel; Table S4). A similar trend was observed for casein kinase 1/2 (CK1/2) alpha, which had previously been associated with embryonic development in mice (Figure 5D, right panel; Table S4) (Dominguez et al., 2011; Gotz and Montenarh, 2017; Lou et al., 2008).

ADME/Tox protein expression profiles of iPSC-derived hepatocytes do not closely resemble PHHs

Drug effects on proteins related to absorption, distribution, metabolism, and excretion, as well as toxicity (ADME/Tox), are critical for the assessment of whether a bioactive compound can be developed as a therapeutic drug. Therefore, we analyzed the proteomic data with a particular emphasis on these proteins (n = 104) (Schroder et al., 2013). Clustering their abundances clearly showed that most proteins were highly expressed only in PHHs (Figure 5E, clusters 5 and 6). Cluster 5 includes GSS, which catalyzes the second step of glutathione synthesis and is an important antioxidant that is used in multiple assays to detect drug toxicity. Clusters 3 and 4 almost exclusively comprise proteins with elevated levels at the MH of 2D differentiated hepatocytes but also contains FOXA1, an important TF for DE development (Lau et al., 2018), early lineage specification, and indicating incomplete maturation (Figure 5F). The latter is also apparent when quantifying the levels of particular proteins (Figure 5F). Although the levels of, e.g., the transporter ABCC1 and GSS are not very different between iPSC-derived hepatocytes and PHHs, other proteins showed 10- to 1,000-fold lower values.

DISCUSSION

The immature properties and incomplete molecular characterization of stem cell-derived hepatocytes hamper their broader application in the assessment of drug toxicity and regenerative medicine. Although previous studies mainly focused on transcriptome analysis (Camp et al., 2017; Ang et al., 2018), few protein markers (Takebe et al., 2017; Vyas et al., 2018; Sekine et al., 2020) and proteomics datasets are available for this biological system (Baxter et al., 2015; Hurrell et al., 2019). Hence we sought to obtain a broader view of hepatocyte differentiation at the protein and PTM levels. These data formed the basis for constructing a detailed and high temporal resolution map of key proteins and pathways involved in the process (Figure 6).

Only a few interesting cases could be highlighted here, and many further global and focused interrogations of the proteomic, phosphoproteomic, and acetylproteomic data can be envisaged. Therefore, all proteomic data are made publicly available for further mining. Rapidly dividing cells, such as stem cells or cancer cells, are tightly associated with the Warburg effect, the phenomenon that cells produce ATP by aerobic glycolysis instead of more energy-efficient oxidative phosphorylation (Liberti and Locasale, 2016; DeBerardinis et al., 2008; Varum et al., 2011). High proliferation rates also come with the need to generate membranes, which is reflected by high rates of fatty acid synthesis (Currie et al., 2013; Rohrig and Schulze, 2016). In line with the above, a key enzyme for *de novo* lipogenesis, FASN, was the most abundant protein in the TkDA3–4 cells and among the top 20 in Ff-I01 cells. On DE specification, we observed a strong drop in the protein levels of FASN, as well as ACACA, the rate-limiting enzyme for *de novo* fatty acid biosynthesis, suggesting that cells shift their metabolic profiles. A similar effect was previously described for FASN during initiation of neural stem cells differentiation (Knobloch et al., 2013). An even more drastic metabolic switch occurred at the transition from HE to IH. High levels of SIRT1 in iPSCs have previously been associated with pluripotency, and diminishing levels of SIRT1 at this stage likely reflect the loss of this cellular characteristic (Lee et al., 2012; Cha et al., 2017; Choi and Mostoslavsky, 2014; Tobita et al., 2016). SIRT2 showed a diametrically opposing expression behavior to SIRT1. High levels of SIRT2 at IH coincided with the observed metabolic switch and suggest the deacetylation of multiple putative lysine residues. The expression changes of both SIRT1 and SIRT2 appear tightly regulated and required for a successful differentiation. The high temporal resolution of the proteomic experiment suggests that the addition of SIRT1 inhibitors at the HE stage could improve the protocol, especially because SIRT1 inhibition has previously been shown to enhance differentiation in the context of neural development and hematopoiesis (Kim et al., 2016; Park et al., 2019a; Hu et al., 2014).

Analysis of the proteome and phosphoproteome also highlighted processes that could not be discerned from the transcriptomic data. For instance, reconstructing kinase-substrate relationships from the phosphoproteomic data allowed us to elucidate the temporal activity of CDK1 during differentiation. Consistent with previous studies (Boward et al., 2016; Neganova et al., 2014) showing high CDK activity in embryonic stem cells, we observed an immediate decrease of substrate phosphorylation on cell differentiation, adding another level of information to the transcriptomic and proteomic data. Previous work has highlighted a complex role of ERK activity in early
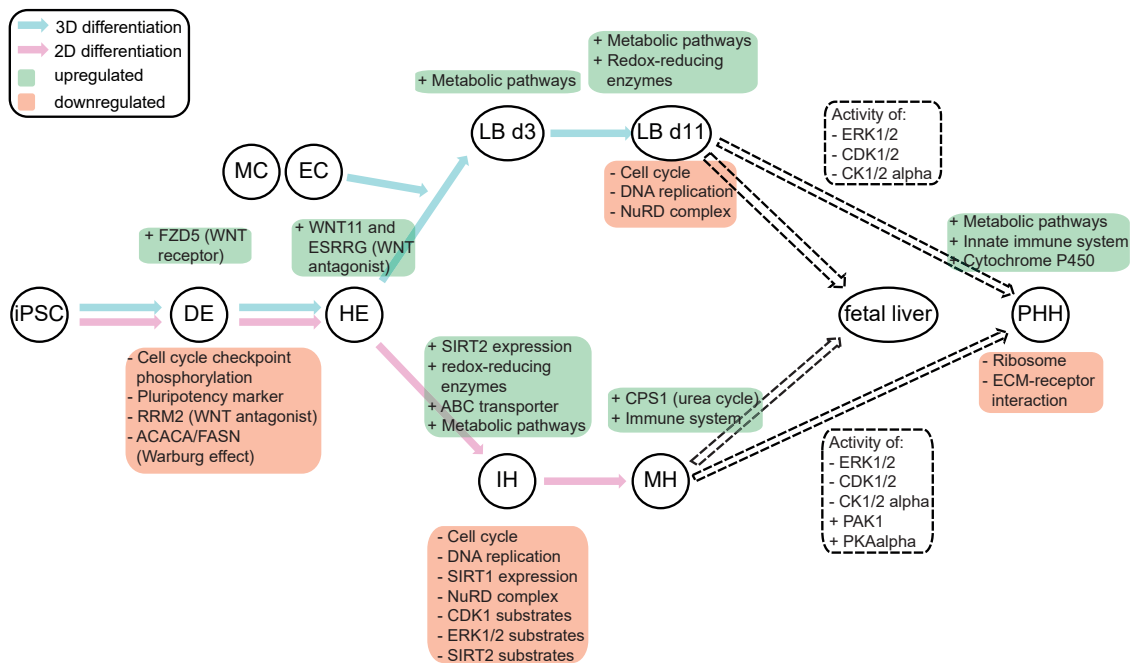
Figure 6. Roadmap of hepatocyte differentiation
Summary of protein and pathways regulated at distinct time points during the process of *in vitro* hepatocyte maturation.

lineage specification (Kunath et al., 2007; Ying et al., 2008; Roode et al., 2012; Piliszek et al., 2017; Yu et al., 2018), but its influence on late hepatocyte maturation had remained inconclusive. Based on our experiments, ERK signaling is not sustained beyond the HE stage, despite increasing protein levels. Although CDK1 and ERK1/2 substrate phosphorylation dropped during *in vitro* differentiation, the comparison with *in vivo* liver samples suggested still higher activity. The same trend was observed for CK1 and CK2 alpha. CK and ERK inhibitors have also been used before to promote differentiation processes, but not in the context of hepatocyte differentiation (Meng et al., 2018; Yu et al., 2018; Jaiswal et al., 2000). Because CK1 inhibitors can enhance the differentiation of pancreatic progenitor cells, which share DE as a common origin with hepatocytes (Zhang et al., 2021), these inhibitors may also work for hepatocyte differentiation.

In summary, our work provides a resource of thousands of proteins and P-sites collected at high temporal resolution that improves our understanding of hepatocyte differentiation. The depicted pathways and proteins specific to discrete differentiation stages add to the functional characterization of the process. Furthermore, our datasets substantiate the reported superiority of 3D models compared with 2D monolayers (Baxter et al., 2015; Chandrasekaran et al., 2017; Branco et al., 2019; Duval et al., 2017; Bratt-Leal et al., 2009) and expose the shortcomings of these systems compared with hepatocytes generated *in vivo*. Still, given the high temporal resolution and proteomic depth of our data, we envision that researchers will find more examples for candidate proteins that characterize a particular cellular state and that may be used to develop improved differentiation protocols in the future.

Limitations of the study

Current MS-based proteomics still require starting materials that are incommensurate of single-cell analysis that may have provided more detailed insights. Limits in sample availability also impact the comprehensiveness of PTM analysis, which scales with the amount of starting material. Although a reasonable coverage of the phosphoproteome could be achieved in this study, it is certainly incomplete and may not include even functionally well-annotated P-sites. This issue is even more pronounced for the acetylproteome in this study. The majority of the presented work is descriptive. Because *in vitro* differentiation experiments are demanding and yield scarce sample quantities, the number of follow-up experiments in this study is limited. Hence several findings and the hypothesis need additional experiments.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cells
  - Primary cells
- METHOD DETAILS
  - 2D hepatocyte differentiation
  - Differentiation of ECs and MCs

- ○ 3D LB generation
- ○ Adult and fetal liver cells
- ○ RNA extraction and qPCR analyis
- ○ Cell lysis, protein digestion, and labeling
- ○ Peptide enrichment, immunoprecipitation, and off-line fractionation
- ○ Mass spectrometry data acquisition
- ○ Peptide and protein database searching
- ○ Proteomics data analysis
- ○ Phosphoproteomics and acetyl-IP data analysis
- ○ scRNA-seq data analysis
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

Conceptualization, J.K., B.T., and B.K.; methodology, J.K., K.S., A.B., M.B., R.Y., A.K., H.T., B.T., J.G.C., and B.K.; software, J.K., P.S., and M.W.; validation, J.K., P.S., M.W., M.B., A.K., and B.K.; formal analysis, J.K., P.S., M.W., and B.K.; investigation, J.K., K.S., A.B., R.Y., and M.B.; resources, J.K., K.S., A.B., M.W., B.T., J.G.C., and B.K.; data curation, J.K., P.S., and M.W.; writing – original draft, J.K. and B.K.; visualization, J.K., P.S., and M.W.; supervision, B.T., J.G.C., and B.K.; project administration, J.K. and B.K.; funding acquisition, B.T. and B.K.

DECLARATION OF INTERESTS

REFERENCES

Ahn, Y., Sims, C., Murray, M.J., Kuhlmann, P.K., Fuentes-Antras, J., Weatherbee, S.D., and Krumlauf, R. (2017). Multiple modes of Lrp4 function in modulation of Wnt/beta-catenin signaling during tooth development. Development 144, 2824–2836.

Alper, C.A., Johnson, A.M., Birtch, A.G., and Moore, F.D. (1969). Human C'3: evidence for the liver as the primary site of synthesis. Science 163, 286–288.

Ang, L.T., Tan, A.K.Y., Autio, M.I., Goh, S.H., Choo, S.H., Lee, K.L., Tan, J., Pan, B., Lee, J.J.H., Lum, J.J., et al. (2018). A roadmap for human liver differentiation from pluripotent stem cells. Cell Rep. 22, 2190–2205.

Bartfeld, S., and Clevers, H. (2017). Stem cell-derived organoids and their application for medical research and patient treatment. J. Mol. Med. 95, 729–738.

Basta, J., and Rauchman, M. (2015). The nucleosome remodeling and deacetylase complex in development and disease. Transl. Res. 165, 36–47.

Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A.P., Cerciello, F., Jacobs, A., Moest, H., Omasits, U., Gundry, R.L., Yoon, C., et al. (2015). A mass spectrometric-derived cell surface protein atlas. PLoS One 10, e0121314.

Baxter, M., Withey, S., Harrison, S., Segeritz, C.P., Zhang, F., Atkinson-Dell, R., Rowe, C., Gerrard, D.T., Sison-Young, R., Jenkins, R., et al. (2015). Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes. J. Hepatol. 62, 581–589.

Berg, T., Rountree, C.B., Lee, L., Estrada, J., Sala, F.G., Choe, A., Veltmaat, J.M., De Langhe, S., Lee, R., Tsukamoto, H., et al. (2007). Fibroblast growth factor 10 is critical for liver growth during embryogenesis and controls hepatoblast survival via beta-catenin activation. Hepatology 46, 1187–1197.

Bisson, J.A., Mills, B., Paul Helt, J.C., Zwaka, T.P., and Cohen, E.D. (2015). Wnt5a and Wnt11 inhibit the canonical Wnt pathway and promote cardiac progenitor development via the Caspase-dependent degradation of AKT. Dev. Biol. 398, 80–96.

Boward, B., Wu, T., and Dalton, S. (2016). Concise review: control of cell fate through cell cycle and pluripotency networks. Stem Cells 34, 1427–1436.

Branco, M.A., Cotovio, J.P., Rodrigues, C.A.V., Vaz, S.H., Fernandes, T.G., Moreira, L.M., Cabral, J.M.S., and Diogo, M.M. (2019). Transcriptomic analysis of 3D cardiac differentiation of human induced pluripotent stem cells reveals faster cardiomyocyte maturation compared to 2D culture. Sci. Rep. 9, 9229.

Bratt-Leal, A.M., Carpenedo, R.L., and Mcdevitt, T.C. (2009). Engineering the embryoid body microenvironment to direct embryonic stem cell differentiation. Biotechnol. Prog. 25, 43–51.

Calvanese, V., Lara, E., Suarez-Alvarez, B., Abu Dawud, R., Vazquez-Chantada, M., Martinez-Chantar, M.L., Embade, N., Lopez-Nieva, P., Horrillo, A., Hmadcha, A., et al. (2010). Sirtuin 1 regulation of developmental genes during differentiation of stem cells. Proc. Natl. Acad. Sci. U S A 107, 13736–13741.

Camp, J.G., Sekine, K., Gerber, T., Loeffler-Wirth, H., Binder, H., Gac, M., Kanton, S., Kageyama, J., Damm, G., Seehofer, D., et al. (2017). Multilineage communication regulates human liver bud development from pluripotency. Nature 546, 533–538.

Cha, Y., Han, M.J., Cha, H.J., Zoldan, J., Burkart, A., Jung, J.H., Jang, Y., Kim, C.H., Jeong, H.C., Kim, B.G., et al. (2017). Metabolic control of primed human pluripotent stem cell fate and function by the miR-200c-SIRT2 axis. Nat. Cell Biol. 19, 445–456.

Challen, G.A., Sun, D., Mayle, A., Jeong, M., Luo, M., Rodriguez, B., Mallaney, C., Celik, H., Yang, L., Xia, Z., et al. (2014). Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. Cell Stem Cell 15, 350–364.

Chandrasekaran, A., Avci, H.X., Ochalek, A., Rosingh, L.N., Molnar, K., Laszlo, L., Bellak, T., Teglasi, A., Pesti, K., Mike, A., et al. (2017). Comparison of 2D and 3D neural induction methods for the generation of neural progenitor cells from human induced pluripotent stem cells. Stem Cell Res. 25, 139–151.

Choi, J.E., and Mostoslavsky, R. (2014). Sirtuins, metabolism, and DNA repair. Curr. Opin. Genet. Dev. 26, 24–32.

Coulonval, K., Kooken, H., and Roger, P.P. (2011). Coupling of T161 and T14 phosphorylations protects cyclin B-CDK1 from premature activation. Mol. Biol. Cell 22, 3971–3985.

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

Cox, J., and Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinform. 13, S12.

Currie, E., Schulze, A., Zechner, R., Walther, T.C., and Farese, R.V., JR. (2013). Cellular fatty acid metabolism and cancer. Cell Metab. 18, 153–161.

Davidson, M.D., Ware, B.R., and Khetani, S.R. (2015). Stem cell-derived liver cells for drug testing and disease modeling. Discov. Med. 19, 349–358.

DeBerardinis, R.J., Lum, J.J., Hatzivassiliou, G., and Thompson, C.B. (2008). The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. Cell Metab. 7, 11–20.

Ding, C., Luo, J., Li, L., Li, S., Yang, L., Pan, H., Liu, Q., Qin, H., Chen, C., and Feng, J. (2016). Gab2 facilitates epithelial-to-mesenchymal transition via the MEK/ERK/MMP signaling in colorectal cancer. J. Exp. Clin. Cancer Res. *35*, 5.

Dominguez, I., Degano, I.R., Chea, K., Cha, J., Toselli, P., and Seldin, D.C. (2011). CK2alpha is essential for embryonic morphogenesis. Mol. Cell Biochem. *356*, 209–216.

Duval, K., Grover, H., Han, L.H., Mou, Y., Pegoraro, A.F., Fredberg, J., and Chen, Z. (2017). Modeling physiological events in 2D vs. 3D cell culture. Physiology (Bethesda) *32*, 266–277.

Estefania, M.M., Ganier, O., Hernandez, P., Schvartzman, J.B., Mechali, M., and Krimer, D.B. (2012). DNA replication fading as proliferating cells advance in their commitment to terminal differentiation. Sci. Rep. *2*, 279.

Ferguson, S.M. (2015). Beyond indigestion: emerging roles for lysosome-based signaling in human disease. Curr. Opin. Cell Biol. *35*, 59–68.

Gotz, C., and Montenarh, M. (2017). Protein kinase CK2 in development and differentiation. Biomed. Rep. *6*, 127–133.

Han, H., Cho, J.W., Lee, S., Yun, A., Kim, H., Bae, D., Yang, S., Kim, C.Y., Lee, M., Kim, E., et al. (2018). TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res. *46*, D380–D386.

Horn, H., Schoof, E.M., Kim, J., Robin, X., Miller, M.L., Diella, F., Palma, A., Cesareni, G., Jensen, L.J., and Linding, R. (2014). KinomeXplorer: an integrated platform for kinome biology studies. Nat. Methods *11*, 603–604.

Hu, B., Guo, Y., Chen, C., Li, Q., Niu, X., Guo, S., Zhang, A., Wang, Y., and Deng, Z. (2014). Repression of SIRT1 promotes the differentiation of mouse induced pluripotent stem cells into neural stem cells. Cel.l Mol. Neurobiol. *34*, 905–912.

Hurrell, T., Segeritz, C.-P., Vallier, L., Lilley, K.S., and Cromarty, A.D. (2019). A proteomic time course through the differentiation of human induced pluripotent stem cells into hepatocyte-like cells. Sci. Rep. *9*, 3270.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol. Cell Biol. *24*, 8862–8871.

Jaiswal, R.K., Jaiswal, N., Bruder, S.P., Mbalaviele, G., Marshak, D.R., and Pittenger, M.F. (2000). Adult human mesenchymal stem cell differentiation to the osteogenic or adipogenic lineage is regulated by mitogen-activated protein kinase. J. Biol. Chem. *275*, 9645–9652.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127.

Jung, J., Zheng, M., Goldfarb, M., and Zaret, K.S. (1999). Initiation of mammalian liver development from endoderm by fibroblast growth factors. Science *284*, 1998–2003.

Kang, M.H., Choi, H., Oshima, M., Cheong, J.H., Kim, S., Lee, J.H., Park, Y.S., Choi, H.S., Kweon, M.N., Pack, C.G., et al. (2018). Estrogen-related receptor gamma functions as a tumor suppressor in gastric cancer. Nat. Commun. *9*, 1920.

Karagiannis, P., Takahashi, K., Saito, M., Yoshida, Y., Okita, K., Watanabe, A., Inoue, H., Yamashita, J.K., Todani, M., Nakagawa, M., et al. (2019). Induced pluripotent stem cells and their use in human models of disease and development. Physiol. Rev. *99*, 79–114.

Kegel, V., Deharde, D., Pfeiffer, E., Zeilinger, K., Seehofer, D., and Damm, G. (2016). Protocol for isolation of primary human hepatocytes and corresponding major populations of non-parenchymal liver cells. J. Vis. Exp *30*, e53069.

Kim, B.S., Lee, C.H., Chang, G.E., Cheong, E., and Shin, I. (2016). A potent and selective small molecule inhibitor of sirtuin 1 promotes differentiation of pluripotent P19 cells into functional neurons. Sci. Rep. *6*, 34324.

Knobloch, M., Braun, S.M., Zurkirchen, L., Von Schoultz, C., Zamboni, N., Arauzo-Bravo, M.J., Kovacs, W.J., Karalay, O., Suter, U., Machado, R.A., et al. (2013). Metabolic control of adult neural stem cell activity by Fasn-dependent lipogenesis. Nature *493*, 226–230.

Kobayashi, T., Piao, W., Takamura, T., Kori, H., Miyachi, H., Kitano, S., Iwamoto, Y., Yamada, M., Imayoshi, I., Shioda, S., et al. (2019). Enhanced lysosomal degradation maintains the quiescent state of neural stem cells. Nat. Commun. *10*, 5446.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Lade, A.G., and Monga, S.P. (2011). Beta-catenin signaling in hepatic development and progenitors: which way does the WNT blow? Dev. Dyn. *240*, 486–500.

Lau, H.H., Ng, N.H.J., Loo, L.S.W., Jasmen, J.B., and Teo, A.K.K. (2018). The molecular functions of hepatocyte nuclear factors - in and beyond the liver. J. Hepatol. *68*, 1033–1048.

Lee, Y.L., Peng, Q., Fong, S.W., Chen, A.C., Lee, K.F., Ng, E.H., Nagy, A., and Yeung, W.S. (2012). Sirtuin 1 facilitates generation of induced pluripotent stem cells from mouse embryonic fibroblasts through the miR-34a and p53 pathways. PLoS One *7*, e45633.

Li, Q., Hutchins, A.P., Chen, Y., Li, S., Shan, Y., Liao, B., Zheng, D., Shi, X., Li, Y., Chan, W.Y., et al. (2017). A sequential EMT-MET mechanism drives the differentiation of human embryonic stem cells towards hepatocytes. Nat. Commun. *8*, 15166.

Liang, R., Arif, T., Kalmykova, S., Kasianov, A., Lin, M., Menon, V., Qiu, J., Bernitz, J.M., Moore, K., Lin, F., et al. (2020). Restraining lysosomal activity preserves hematopoietic stem cell quiescence and potency. Cell Stem Cell *26*, 359–376.e7.

Liberti, M.V., and Locasale, J.W. (2016). The Warburg effect: how does it benefit cancer cells? Trends Biochem. Sci. *41*, 211–218.

Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. Cell *165*, 535–550.

Lou, D.Y., Dominguez, I., Toselli, P., Landesman-Bollag, E., O'brien, C., and Seldin, D.C. (2008). The alpha catalytic subunit of protein kinase CK2 is required for mouse embryonic development. Mol. Cell Biol. *28*, 131–139.

McLean, A.B., D'amour, K.A., Jones, K.L., Krishnamoorthy, M., Kulik, M.J., Reynolds, D.M., Sheppard, A.M., Liu, H., Xu, Y., Baetge, E.E., and Dalton, S. (2007). Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. Stem Cells *25*, 29–38.

Medvedeva, Y.A., Lennartsson, A., Ehsani, R., Kulakovskiy, I.V., Vorontsov, I.E., Panahandeh, P., Khimulya, G., Kasukawa, T., Consortium, F., and Drablos, F. (2015). EpiFactors: a comprehensive database of human epigenetic factors and complexes. Database (Oxford) *2015*, bav067.

Meng, Y., Ren, Z., Xu, F., Zhou, X., Song, C., Wang, V.Y., Liu, W., Lu, L., Thomson, J.A., and Chen, G. (2018). Nicotinamide promotes cell survival and differentiation as kinase inhibitor in human pluripotent stem cells. Stem Cell Rep. *11*, 1347–1356.

Monteleon, C.L., Agnihotri, T., Dahal, A., Liu, M., Rebecca, V.W., Beatty, G.L., Amaravadi, R.K., and Ridky, T.W. (2018). Lysosomes support the degradation, signaling, and mitochondrial metabolism necessary for human epidermal differentiation. J. Invest. Dermatol. *138*, 1945–1954.

Neganova, I., Tilgner, K., Buskin, A., Paraskevopoulou, I., Atkinson, S.P., Peberdy, D., Passos, J.F., and Lako, M. (2014). CDK1 plays an important role in the maintenance of pluripotency and genomic stability in human pluripotent stem cells. Cell Death Dis. *5*, e1508.

Park, J.A., Park, S., Park, W.Y., Han, M.K., and Lee, Y. (2019a). Splitomicin, a SIRT1 inhibitor, enhances hematopoietic differentiation of mouse embryonic stem cells. Int. J. Stem Cells *12*, 21–30.

Park, M.J., D'alecy, L.G., Anderson, M.A., Basrur, V., Feng, Y., Brady, G.F., Kim, D.I., Wu, J., Nesvizhskii, A.I., Lahann, J., et al. (2019b). Constitutive release of CPS1 in bile and its role as a protective cytokine during acute liver injury. Proc. Natl. Acad. Sci. U S A *116*, 9125–9134.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., GRISS, J., Mayer, G., Eisenacher, M., et al. (2019).

The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.

Piliszek, A., Madeja, Z.E., and Plusa, B. (2017). Suppression of ERK signalling abolishes primitive endoderm formation but does not promote pluripotency in rabbit embryo. Development *144*, 3719–3730.

Qu, X., Yu, B., Zhu, M., Li, X., Ma, L., Liu, C., Zhang, Y., and Cheng, Z. (2021). Sinomenine inhibits the growth of ovarian cancer cells through the suppression of mitosis by down-regulating the expression and the activity of CDK1. Onco Targets Ther. *14*, 823–834.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., et al. (2012). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. Cell Stem Cell *10*, 583–594.

Rigbolt, K.T., Prokhorova, T.A., Akimov, V., Henningsen, J., Johansen, P.T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J.V., and Blagoev, B. (2011). System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci. Signal *4*, rs3.

Rohrig, F., and Schulze, A. (2016). The multifaceted roles of fatty acid synthesis in cancer. Nat. Rev. Cancer *16*, 732–749.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev. Biol. *361*, 358–363.

Rowe, C., Gerrard, D.T., Jenkins, R., Berry, A., Durkin, K., Sundstrom, L., Goldring, C.E., Park, B.K., Kitteringham, N.R., Hanley, K.P., and Hanley, N.A. (2013). Proteome-wide analyses of human hepatocytes during differentiation and dedifferentiation. Hepatology *58*, 799–809.

Ruprecht, B., Koch, H., Domasinska, P., Frejno, M., Kuster, B., and Lemeer, S. (2017a). Optimized enrichment of phosphoproteomes by Fe-IMAC column chromatography. Methods Mol. Biol. *1550*, 47–60.

Ruprecht, B., Zecha, J., Zolg, D.P., and Kuster, B. (2017b). High pH reversed-phase micro-columns for simple, sensitive, and efficient fractionation of proteome and (TMT labeled) phosphoproteome digests. Methods Mol. Biol. *1550*, 83–98.

Sato, A., Yamamoto, H., Sakane, H., Koyama, H., and Kikuchi, A. (2010). Wnt5a regulates distinct signalling pathways by binding to Frizzled2. EMBO J. *29*, 41–54.

Schroder, A., Klein, K., Winter, S., Schwab, M., Bonin, M., Zell, A., and Zanger, U.M. (2013). Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. Pharmacogenomics J. *13*, 12–20.

Sekine, K., Ogawa, S., Tsuzuki, S., Kobayashi, T., Ikeda, K., Nakanishi, N., Takeuchi, K., Kanai, E., Otake, Y., Okamoto, S., et al. (2020). Generation of human induced pluripotent stem cell-derived liver buds with chemically defined and animal origin-free media. Sci. Rep. *10*, 17937.

Shen, C.N., Slack, J.M., and Tosh, D. (2000). Molecular basis of transdifferentiation of pancreas to liver. Nat. Cell Biol. *2*, 879–887.

Si, X., Chen, W., Guo, X., Chen, L., Wang, G., Xu, Y., and Kang, J. (2013). Activation of GSK3beta by Sirt2 is required for early lineage commitment of mouse embryonic stem cell. PLoS One *8*, e76699.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.R., Ueno, Y., Zheng, Y.W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature *499*, 481–484.

Takebe, T., Sekine, K., Kimura, M., Yoshizawa, E., Ayano, S., Koido, M., Funayama, S., Nakanishi, N., Hisai, T., Kobayashi, T., et al. (2017). Massive and reproducible production of liver buds entirely from human pluripotent stem cells. Cell Rep. *21*, 2661–2670.

Takebe, T., Zhang, R.R., Koike, H., Kimura, M., Yoshizawa, E., Enomura, M., Koike, N., Sekine, K., and Taniguchi, H. (2014). Generation of a vascularized and functional human liver from an iPSC-derived organ bud transplant. Nat. Protoc. *9*, 396–409.

Tobita, T., Guzman-Lepe, J., Takeishi, K., Nakao, T., Wang, Y., Meng, F., Deng, C.X., Collin De L'hortet, A., and Soto-Gutierrez, A. (2016). SIRT1 disruption in human fetal hepatocytes leads to increased accumulation of glucose and lipids. PLoS One *11*, e0149344.

Toivonen, S., Lundin, K., Balboa, D., Ustinov, J., Tamminen, K., Palgi, J., Trokovic, R., Tuuri, T., and Otonkoski, T. (2013). Activin A and Wnt-dependent specification of human definitive endoderm cells. Exp. Cell Res. *319*, 2535–2544.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Van Hoof, D., Munoz, J., Braam, S.R., Pinkse, M.W., Linding, R., Heck, A.J., Mummery, C.L., and Krijgsveld, J. (2009). Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell *5*, 214–226.

Varum, S., Rodrigues, A.S., Moura, M.B., Momcilovic, O., Easley, C.A.T., Ramalho-Santos, J., Van Houten, B., and SCHATTEN, G. (2011). Energy metabolism in human pluripotent stem cells and their differentiated counterparts. PLoS One *6*, e20914.

Vasiliou, V., Vasiliou, K., and Nebert, D.W. (2009). Human ATP-binding cassette (ABC) transporter family. Hum. Genom. *3*, 281–290.

Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M.B., Zinner, M., Olivieri, D., Vabres, P., Duplomb-Jego, L., De Bont, E., et al. (2019). Lysosomal signaling licenses embryonic stem cell differentiation via inactivation of Tfe3. Cell Stem Cell *24*, 257–270.e8.

Vyas, D., Baptista, P.M., Brovold, M., Moran, E., Gaston, B., Booth, C., Samuel, M., Atala, A., and Soker, S. (2018). Self-assembled liver organoids recapitulate hepatobiliary organogenesis in vitro. Hepatology *67*, 750–761.

Wang, D., Eraslan, B., Wieland, T., Hallstrom, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. *15*, e8503.

Westmacott, A., Burke, Z.D., Oliver, G., Slack, J.M., and Tosh, D. (2006). C/EBPalpha and C/EBPbeta are markers of early liver development. Int. J. Dev. Biol. *50*, 653–657.

Wiese, M., and Stefan, S.M. (2019). The A-B-C of small-molecule ABC transport protein modulators: from inhibition to activation-a case study of multidrug resistance-associated protein 1 (ABCC1). Med. Res. Rev. *39*, 2031–2081.

Xu, Y., Li, F., Lv, L., Li, T., Zhou, X., Deng, C.X., Guan, K.L., Lei, Q.Y., and Xiong, Y. (2014). Oxidative stress activates SIRT2 to deacetylate and stimulate phosphoglycerate mutase. Cancer Res. *74*, 3630–3642.

Yang, P., Humphrey, S.J., Cinghu, S., Pathania, R., Oldfield, A.J., Kumar, D., Perera, D., Yang, J.Y.H., James, D.E., Mann, M., and Jothi, R. (2019). Multiomic profiling reveals dynamics of the phased progression of pluripotency. Cell Syst. *8*, 427–445.e10.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Yu, G., and He, Q.Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol. Biosyst. *12*, 477–479.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Yu, P., Petzoldt, S., Wilhelm, M., Zolg, D.P., Zheng, R., Sun, X., Liu, X., Schneider, G., Huhmer, A., and Kuster, B. (2017). Trimodal mixed mode chromatography that enables efficient offline two-dimensional peptide fractionation for proteome analysis. Anal Chem. *89*, 8884–8891.

Yu, Y., Wang, X., Zhang, X., Zhai, Y., Lu, X., Ma, H., Zhu, K., Zhao, T., Jiao, J., Zhao, Z.A., and Li, L. (2018). ERK inhibition promotes neuroectodermal

precursor commitment by blocking self-renewal and primitive streak formation of the epiblast. Stem Cell Res. Ther. *9*, 2.

Zecha, J., Satpathy, S., Kanashova, T., Avanessian, S.C., Kane, M.H., Clauser, K.R., Mertins, P., Carr, S.A., and Kuster, B. (2019). TMT labeling for the masses: a robust and cost-efficient, in-solution labeling approach. Mol. Cell Proteomics *18*, 1468–1478.

Zhang, Y., Xu, J., Ren, Z., Meng, Y., Liu, W., Lu, L., Zhou, Z., and Chen, G. (2021). Nicotinamide promotes pancreatic differentiation through the dual inhibition of CK1 and ROCK kinases in human embryonic stem cells. Stem Cell Res. Ther. *12*, 362.

Zhou, Z., Xu, M.J., and Gao, B. (2016). Hepatocytes: a key cell type for innate immunity. Cell Mol. Immunol. *13*, 301–315.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Fetal liver	UK HDBR resource	N/A
PHH	University Clinic of Leipzig	N/A
Chemicals, peptides, and recombinant proteins		
mTeSR1	StemCell Technologies	Cat# 85850
Laminin 511-E8	iMatrix-511, Nippi	Cat# T304
Accutase®	Sigma-Aldrich	Cat# A6964
RPMI 1640	Gibco™	Cat# 21875034
B-27™	Gibco™	Cat# 17504044
Wnt3a	R&D Systems	Cat# 5036-WNP
Activin A	Sigma-Aldrich	Cat# A4941
Activin A	Ajinomoto	Cat# AMS-ACTA-50
ROCK Inhibitor Y-27632	Fujifilm Wako Pure Chemical	Cat# 259-00613; CAS: 331752-47-7
KnockOut™-DMEM	Gibco™	Cat# 10829018
DMSO	Sigma-Aldrich	Cat# D2438; CAS: 67-68-5
KnockOut™ Serum Replacement	Gibco™	Cat# 10828010
GlutaMAX™	Gibco™	Cat# 35050061
MEM Non-Essential Amino Acids	Gibco™	Cat# 11140050
2-Mercaptoethanol	Gibco™	Cat# 31350010
HCM™ Hepatocyte Culture Medium BulletKit ™	Lonza	Cat# CC-3198
Hepatocyte growth factor	R&D Systems	Cat# 294-HG-005
Oncostatin M	R&D Systems	Cat# 295-OM
StemFit™ For Differentiation	Ajinomoto	Cat# AS401
CHIR99021	Cayman Chemical	Cat# 13122; CAS: 252917-06-9
Sodium butyrate	Sigma-Aldrich	Cat# B5887; CAS: 156-54-7
StemFit™ Basic03	Ajinomoto	Cat# ASB03
L-glutamine	Gibco™	Cat# G7513; CAS: 56-85-9
Dexamethasone	Sigma-Aldrich	Cat# D2915; CAS: 50-02-2
DMEM/F-12	Gibco™	Cat# 11320074
BMP4	R&D Systems	Cat# 314-BP
StemPro-34 SFM medium	Gibco™	Cat# 10639011
VEGF	Gibco™	Cat# PHC9394
Forskolin	Cayman Chemical	Cat# 11018
PDGF-BB	Peprotech	Cat# 100-14B
bFGF	Fujifilm Wako Pure Chemical	Cat# 064-04541
TrypLE™	Gibco™	Cat# 12563029
BD Matrigel™	BD Bioscience	Cat# 356234
DMEM	Fujifilm Wako Pure Chemical	Cat# 043-30085
KBM-VEC1 basal medium	Fujifilm Wako Pure Chemical	Cat# 636-41045
cOmplete™ EDTA-free protease inhibitor	Roche	Cat# 11873580001
Phosphatase inhibitor cocktail 1	Sigma-Aldrich	Cat# P2850
Phosphatase inhibitor cocktail 2	Sigma-Aldrich	Cat# P5726
Phosphatase inhibitor cocktail 3	Sigma-Aldrich	Cat# P0044
Trypsin	Roche	Cat# 11047841001
TMT-6plex	Thermo Scientific™	Cat# 90068
TMT-11plex	Thermo Scientific™	Cat# A34808

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hydroxylamine	Sigma-Aldrich	Cat# 438227
Ammonium hydroxide solution	Sigma-Aldrich	Cat# 221228; CAS: 1336-21-6
Ammonium formate	Sigma-Aldrich	Cat# 770221; CAS: 540-69-2
PTMScan® IAP	Cell Signaling Technologies	Cat# 9993S
Ammonium acetate	Sigma-Aldrich	Cat# 73594, CAS: 631-61-8
Ammonium bicarbonate	Sigma-Aldrich	Cat# 09830; CAS: 1066-33-7
Critical commercial assays		
Pierce™ Coomassie Plus (Bradford) Assay Kit	Thermo Scientific™	Cat# 23236
PTMScan® Acetyl-Lysine Motif Kit	Cell Signaling Technology	Cat# 13416
RNeasy Kit	QIAGEN	Cat# 74104
iScript™ cDNA Synthesis Kit	BioRad	Cat# 1708890
SensiMix™ SYBR® No-ROX Kit	Bioline	Cat# QT650-05
Deposited data		
Protein mass spectrometry data	This paper (Pride)	PXD027439
scRNA-seq data	Camp et al., 2017 (Gene Expression Omnibus)	GSE81252
Experimental models: Cell lines		
TkDA3-4	Takebe et al., 2013	N/A
Ff-I01	Takebe et al., 2017	N/A
Oligonucleotides		
A complete list of qPCR primers can be found in Table S6	This paper	Roche UniversalProbe library system
Software and algorithms		
Adobe Illustrator	Adobe	https://www.adobe.com/
Maxquant software	Cox and Mann, 2008	Version 1.6.2.3: https://www.maxquant.org/
GraphPad Prism 5	GraphPad Software, San Diego, California USA	https://www.graphpad.com/
Perseus	Tyanova et al., 2016	Version 1.6.2.3: https://maxquant.net/perseus/
R	The R Project	Version 4.0.2: https://www.r-project.org/
R Package "sva"	Johnson et al., 2007	Version 3.30.1
R Package "factoextra"		Version 1.0.7: https://cloud.r-project.org/package=factoextra
R Package "ggplot2"		Version 3.3.2: https://cloud.r-project.org/package=ggplot2
R Package "clusterProfiler"	Yu et al., 2012	Version 3.16.0: https://bioconductor.org/packages/clusterProfiler/
TRRUST (Version 2)	Han et al., 2018	https://www.grnpedia.org/trrust/
Networkin	Horn et al., 2014	http://www.networkin.info/
Other		
C18 Extraction Disks	3M™ Empore™	Cat# 11913614
SepPak column	Waters Corps.	Cat# WAT054960
ProPac™ IMAC-10 4x50mm	Thermo Fisher Scientific	Cat# TFS-063276
Acclaim AmG C18 column	Thermo Fisher Scientific	Cat# 088753
Dionex Ultimate 3000 HPLC system	Thermo Fisher Scientific	N/A
XBridge BEH130, 3.5 μm, 2.1x150mm	Waters Corps.	Cat# WT186003565
Fusion Lumos Tribrid mass spectrometer	Thermo Fisher Scientific	Cat# IQLAAEGAAPFADBMBHQ

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Aekta HPLC system	GE Healthcare Life Sciences	N/A
scRNA seq data	Camp et al., 2017	GEO: GSE81252
Human reference proteome	UniProt	UP000005640 (Downloaded 22.07.2013): https://www.uniprot.org/ proteomes/UP000005640

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bernhard Kuster (kuster@tum.de).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository. The accession number is listed in the key resources table. This paper analyzes existing, publicly available scRNA-seq data. The accession number for the dataset is listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cells
Both iPSC lines were kindly provided by The University of Tokyo and maintained on laminin 511-E8 (iMatrix-511, Nippi)-coated dishes under 5% CO2 and 37°C. TkDA3-4 iPSC line (Takebe et al., 2013) was cultivated in standard feeder-free conditions in mTeSR1 (StemCell Technologies). Ff-I01 iPSC line (Takebe et al., 2017) was cultivated in StemFit Basic03 (Ajinomoto) supplemented with 80 ng/mL bFGF (Fujifilm Wako Pure Chemical).

Primary cells
Adult human liver samples were obtained from the Universitätsklinikum Leipzig with the donor''s informed consent. Adult human liver tissue was obtained from the Universitätsklinikum Leipzig with ethical approval from the Ethics Commision of the Medical Faculty of the University of Leipzig (approval number 177/16-1k). Primary human hepatocytes (PHH) were isolated from healthy tissue that remained resected human liver of one patient as described previously (Kegel et al., 2016).

The experiments with human fetal liver were performed with ethical approval from the Ethics Commision of the Medical Faculty of the University of Leipzig (approval number 332/16-ek). The human fetal liver samples were provided by the Joint MRC/Wellcome Trust HDBR from gestation week 16 and 17.

METHOD DETAILS

2D hepatocyte differentiation
2D hepatocytes were generated from two different human-derived iPSC lines, TkDA3-4 (Takebe et al., 2013) and Ff-I01 (Takebe et al., 2017), respectively. Cell lines were confirmed to be mycoplasma negative before differentiations were started. TkDA3-4 cells were cultivated in standard feeder-free conditions in mTeSR1 (StemCell Technologies) on laminin 511-E8 (iMatrix-511, Nippi)-coated dishes and were differentiated like described previously (Camp et al., 2017). Briefly, TkDA3-4 cells were dissociated using Accutase® (Sigma-Aldrich) and seeded in RPMI 1640 (Gibco™) medium with 1% B-27™ (Gibco™), 50 ng/mL Wnt3a (R&D Systems), and 100 ng/mL activin A (Sigma-Aldrich) on laminin 511-E8 (iMatrix-511, Nippi)-coated dishes. For the initial seeding, 10 μM ROCK Inhibitor Y-27632 (Fujifilm Wako Pure Chemical) was supplemented. Between day 6 and day 13, cells were cultivated in KnockOut™-DMEM (Gibco™) with 1% (vol/vol) DMSO (Sigma-Aldrich), 20% (vol/vol) KnockOut™ Serum Replacement, 1 mM GlutaMAX™, 1% (vol/vol) non-essential amino acids, and 0.1 mM 2-mercaptoethanol (all Gibco™). Lastly, cells were cultured in hepatocyte culture medium (Lonza) without EGF and supplemented with 20 ng/mL hepatocyte growth factor and 20 ng/mL oncostatinM (both R&D Systems) until day 21. The medium was exchanged daily during the differentiation process and the cell morphology was

monitored by microscopy. At day 0 (iPSC), day 6 (DE), day 8 (HE), day 13 (IH), and day 21 (MH) cells were harvested and subsequently used for RNA extraction and MS sample preparation. For the higher resolution experiment, samples were taken every 12 h between HE and IH. As shown by Camp et al., close to 100% of the iPSCs developed into DE (Camp et al., 2017). While the subsequent differentiation towards HE showed some heterogeneity, almost 90% of IH cells expressed hepatoblast-like genes such as AFP, TTR, and RBP4.

Hepatocytes derived from Ff-I01 cells were generated like described previously (Sekine et al., 2020). Briefly, cells were cultivated in StemFit[TM] Basic03 (Ajinomoto) with 80 ng/mL bFGF (Fujifilm Wako Pure Chemical) medium and dissociated for differentiation using Accutase® (Sigma-Aldrich). Dishes were coated with laminin 511-E8 (iMatrix-511, Nippi) and cells were seeded in the presence of ROCK inhibitor Y-27632 (Fujifilm Wako Pure Chemical). For the first 6 days of differentiation, cells were cultivated in RPMI 1640 (GIBCO[TM]) supplemented with 20% StemFit[TM] For Differentiation and 100 ng/mL activin A (both Ajinomoto). Additionally, 2 μM CHIR99021 (Cayman Chemical) was added for the first 3 days and 0.5 mM sodium butyrate (Sigma Aldrich) was supplemented from day 1 to day 4 resulting in almost 100% of the cells expressing the DE marker CXCR4 (Sekine et al., 2020). From day 6 to day 13, cells were cultured in StemFit[TM] Basic03 (Ajinomoto) medium supplemented with 1% DMSO (Sigma-Aldrich), 0.1 mM 2-mercaptoethanol, 0.5% L-glutamine, and 1% non-essential amino acids (all Gibco[TM]). The medium was exchanged daily during the first 13 days of differentiation. For the final 8 days, cells were cultured in DMEM medium (GIBCO[TM]) supplemented with 5% StemFit[TM] For Differentiation (Ajinomoto) and 0.1 μM dexamethasone (Sigma-Aldrich) with medium exchange every second day.

Differentiation of ECs and MCs

ECs and MCs were differentiated from iPSC as reported previously (Takebe et al., 2017). Briefly, for EC generation, human iPSCs were dissociated with Accutase® (Sigma-Aldrich) and plated on laminin 511-E8 (iMatrix-511, Nippi)-coated dishes with the mesoderm induction medium consisting of DMEM/F-12 (1:1 mixture) medium supplemented with 1% GlutaMAX[TM], 1% B-27[TM] (all Gibco[TM]), 8 μM CHIR99021 (Cayman Chemical), and 25 ng/mL BMP4 (R&D Systems). After additional 4 days, the mesoderm induction medium was replaced with EC Induction Medium consisting of StemPro-34 SFM medium supplemented with 200 ng/mL VEGF (both Gibco[TM]) and 2 μM forskolin (Cayman Chemical). The medium was exchanged every second day until day 10.

For MC/STM differentiation, human iPSCs were dissociated with Accutase® (Sigma-Aldrich) and plated on laminin 511-E8 (iMatrix-511, Nippi)-coated dishes with mesoderm induction medium consisting of DMEM/F-12 (1:1 mixture) medium supplemented with 1% GlutaMAX[TM], 1% B-27[TM] (all Gibco[TM]), 8 μM CHIR99021 (Cayman Chemical), and 25 ng/mL BMP4 (R&D Systems) for 4 days. Then, cells were cultured with LPM induction medium consisting of DMEM/F-12 supplemented with 1% B-27[TM], 10 ng/mL PDGFBB, 2 ng/mL activin A, and 1% GlutaMAX[TM] for 2 days. Afterwards, cells were cultured with STM induction medium consisting of DMEM/F-12, 1% B-27[TM], 10 ng/mL bFGF, 12 ng/mL BMP4, and 1% GlutaMAX[TM].

3D LB generation

LBs were generated like described previously (Takebe et al., 2014). Briefly, HE, EC, and MC/STM cells were derived from iPSCs, dissociated with TrypLE[TM] (Gibco[TM]) and plated at a ratio of 10:7:2 onto a Matrigel (BD Bioscience)-coated dish. Cells were cultured in 1:1 mixture of DMEM (GIBCO[TM]) and KBM-VEC1 basal medium (Fujifilm Wako Pure Chemical) supplemented with 2.5% StemFit[TM] For Differentiation (Ajinomoto) and 0.1 μM dexamethasone (Sigma-Aldrich). For initial seeding, 10 μM of the ROCK inhibitor Y-27632 was added and thereafter half of the medium was exchanged daily.

Adult and fetal liver cells

Adult human liver samples were obtained from the Universitätsklinikum Leipzig with the donor''s informed consent. Primary human hepatocytes (PHH) were isolated as described previously (Kegel et al., 2016). Briefly, liver tissues were dissociated using a two-step EGTA/collagenase P perfusion incubation and the PHH cells were separated from the non-parenchymal cells by centrifugation at 50 xg.

The human fetal liver samples were provided by the Joint MRC/Wellcome Trust HDBR with ethical approval. In this work, human fetal liver samples from gestation week 16 and 17 were used for the proteomics study.

RNA extraction and qPCR analyis

For qPCR analysis, RNA from the harvested cells was isolated with the RNeasy Kit (QIAGEN) by following the manufacturer''s manual. The RNA concentration was measured via Nanodrop2000 and 100–200 ng of total RNA were reverse transcribed with the iScript[TM] cDNA Synthesis Kit (Bio-Rad). 500 ng of cDNA were further used for each qPCR reaction with the SensiMix SYBR Kit (Bioline) and QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) according to manufacturer's instructions. First, the data was normalized to the housekeeping gene GAPDH (TkDA3-4 cells) or 18 S RNA (Ff-I01 cells), then the averaged $2^{(-\Delta CT)}$ was calculated for each time point and further z-scored for heatmap visualizations. A list of all primers can be found in Table S5.

Cell lysis, protein digestion, and labeling

Cells were harvested/isolated and washed twice with PBS (w/o CaCl$_2$/MgCl$_2$, Sigma-Aldrich). The cell pellet was resuspended in lysis buffer containing 8 M urea (Sigma-Aldrich), 40 mM Tris/HCl (pH 7.6, Roche), 1x EDTA-free protease inhibitor (cOmplete[TM], Roche), and 1x phosphatase inhibitor mix (prepared in-house according to the Phosphatase Inhibitor Cocktail 1, 2, and 3 from

Sigma-Aldrich). Samples were frozen at −80°C and thawed on ice before they were pelleted at 20.000 xg at 4°C for 20 min. The protein concentration was measured via Pierce™ Coomassie (Bradford) and 100 μg (for TkDA3-4 differentiations) or 70 μg (for the Ff-I01 differentiation) were further used for digesting. For this, samples were reduced with 10 mM DTT for 45 min at 37°C and alkylated with 55 mM CAA for 30 min at RT in the dark. Next, 40 mM Tris was used to decrease the urea concentration below 1.6 M. The digestion was started by adding 1:100 (enzyme:protein) trypsin (Roche) and incubating at 37°C for 3 h and 700 rpm on a thermoshaker. After the pre-incubation, more trypsin was added at a 1:100 (enzyme:protein) ratio and incubated overnight. After digestion, samples were acidified to a pH of 2–3 with formic acid (FA, Sigma-Aldrich) and used for StageTip desalting like described previously (Rappsilber et al., 2007). The desalted samples were frozen and dried using a Speed-Vac. Next, the peptides were labeled with TMT (Thermo Scientific) like described previously (Zecha et al., 2019). Briefly, peptides were reconstituted in 20 μL of 50 mM HEPES (pH 8.5), 5 μL of 11.6 mM TMT reagent was added and incubated at RT shaking at 400 rpm. After 1 h, the labeling reaction was stopped using 2 μL of 5% hydroxylamine (Sigma-Aldrich). Next, all channels were pooled and the reaction tubes were rinsed with 20 μL of 10% FA in 10% ACN and also combined to the pooled samples. The labeled peptides were frozen and dried using a Speed-Vac. The labeled peptides were further reconstituted in 0.1% FA and loaded on pre-equilibrated 50 mg SepPak columns (Water Corp.). Unspecific binders were washed away with 0.1% FA, peptides were eluted with 0.1% FA in 50% ACN, and dried.

Peptide enrichment, immunoprecipitation, and off-line fractionation

Next, phosphopeptides were enriched as described earlier (Ruprecht et al., 2017a). In brief, desalted samples were reconstituted in 0.07% TFA with 30% ACN and loaded onto a Fe^{3+}-IMAC column (ProPac™ IMAC-10 4 × 50 mm, Thermo Fisher Scientific) coupled to an AEKTA HPLC system (GE Healthcare Life Sciences). The unbound peptides were collected and further used as full proteome (FP). The bound phosphopeptides were eluted with 0.315% ammonium hydroxyde and fractionated as described previously (Ruprecht et al., 2017b). In Brief, dried phosphopeptides were resuspended in 25 mM ammonium formate, loaded onto StageTips (5 disks, diameter 1.5 mm C18 material, 3 M Empore) and eluted with 5%, 10%, 15%, 17.5%, and 50% ACN in 25 mM ammonium formate. The 5% and 50% fractions as well as the 17.5% and flowthrough were combined. All fractions were frozen and dried before data acquisition.

The TkDA3-4 FP was further used for acetyl-IP according to the manufacturer's protocol with some modifications. For this, the dried FP samples were reconstituted in PTMScan® IAP buffer and incubated at 4°C for 1 h with PTMScan® Acetyl-Lysine Motif (both Cell Signaling Technology) antibody beads. The unbound FP was dried and used for further offline fractionation. Acetylated peptides were eluted with 0.15% TFA and desalted via StageTips (3 disks, diameter 1.5 mm C18 material, 3 M Empore) (Rappsilber et al., 2007). Desalted samples were further dried and used for MS measurements.

The FP fractions were further deep fractionated via Trinity fractionation (only TkDA3-4 cells) or hPH reversed phase as described previously (Yu et al., 2017). For Trinity fractionation, samples were reconstituted in 10 mM ammonium acetate (in water, pH 4.7) and loaded onto an Acclaim AmG C18 column (2.1 × 150 mm, Thermo Scientific) connected to a Dionex Ultimate 3000 HPLC system (Thermo Fisher). Peptides were eluted by increasing the concentration of the elution buffer (10 mM ammonium acetate in 95% ACN, pH 5.4) and pooled to 32 fractions. For hPH fractionation, samples were reconstituted in 25 mM ammonium bicarbonate (pH 8) and loaded onto a C18 column (XBridge BEH130, 3.5 μm, 2.1 × 150 mm, Waters Corp.). Peptides were eluted with increasing ACN concentration to 96 fractions and further pooled to 48 fractions.

Mass spectrometry data acquisition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PRIDE: PXD027439. All samples were measured in data-dependent acquisition mode using a nanoflow LC-MS/MS by coupling a Dionex Ultimate 3000 UHPLC + system to a Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific). Peptides were loaded onto an in-house packed trap column (75 μm × 2 cm, 5 μm C18 resin; Reprosil PUR AQ, Dr. Maisch) and separated using an analytical column (75 μm × 40 cm, packed in-house with 3 μm C18 resin; Reprosil PUR AQ) with a flow rate of 300 nL/min. The following MS method was used and adapted depending on the sample type and labelling reagent.

The TMT6-plex labeled fullproteome was eluted with a linear 50 min gradient from 8% to 34% LC buffer B (0.1% FA, 5% DMSO in ACN) in LC buffer A (0.1% FA, 5% DMSO in MS-grade water). The MS1 spectra were acquired in positive mode in the orbitrap with 60,000 resolution and a scan range from 360-1300 m/z (automatic gain control (AGC) target of 4e5 charges, maximum injection time (maxIT) of 20 ms). The isolated precursors were fragmented via CID (collision induced dissociation) and the resulting MS2 spectra were recorded with 15,000 resolution in the orbitrap, an AGC target of 5e4 charges, and a maxIT of 22 ms. For the subsequent SPS-MS3 scans, 10 MS2 fragments were selected simultaneously and further fragmented via HCD (higher-energy collisional dissociation). The resulting ions were recorded in the orbitrap with 15,000 resolution, an AGC target of 1e5 charges, and a maxIT of 50 ms. For the TMT11-plex labeled FP, the same LC gradient was used and only the following changes were made. The MS2 spectra were recorded in the iontrap with rapid mode, an AGC target of 2e4 charges, and a maxIT time of 60 ms. The MS3 spectra were acquired with 50,000 resolution, an AGC target of 1.2e5 charges, and a maxIT of 120 ms.

The IMAC-enriched phosphopeptides were eluted with a linear 80 min gradient from 4% to 32% LC buffer B and the MS1 spectra were subsequently recorded in the orbitrap with 60,000 resolution, an AGC target of 4e5 charges, and a maxIT of 20 ms (for TMT6 samples) or 50 ms (for TMT11 samples), respectively. For the following MS2 spectra, precursor ions were fragmented via CID and

recorded in the orbitrap with 15,000 resolution, an AGC target of 5e4 charges, and a maximum IT of 22 ms. The SPS-MS3 scan was acquired in the orbitrap with 15,000 resolution, an AGC target of 1e5 charges, and a maximum IT of 50 ms for the TMT6-plex labeled phosphopeptides. For the TMT11-plex labeled samples, a resolution of 50,000, an AGC target of 1.2e5 charges, and a maxIT of 120 ms was used.

Samples from the acetyl-IP were eluted with a linear 100 min gradient from 6% to 34% LC buffer B. MS1 spectra were acquired with 60,000 resolution in the orbitrap, an AGC target of 4e5 charges, and a maxIT of 20 ms. For the subsequent MS2 spectra, precursor ions were fragmented via CID and recorded in the orbitrap with 15,000 resolution, an AGC target of 1e5 charges, and a maxIT of 200 ms. For the following SPS-MS3 scan, 10 precursors were selected, fragmented via HCD, and read out in the orbitrap with 15,000 resolution (AGC target of 1.2e5 charges and a maxIT of 300 ms).

Peptide and protein database searching

Raw files were searched with MaxQuant software (Cox and Mann, 2008) (version [1.6.2.3]) against the UniProtKB human reference list (downloaded 22.07.2013) with default settings unless stated otherwise. Trypsin was used as protease and up to two missed cleavages were allowed. Carbamidomethylation of cysteine was set as a fixed modification and oxidation of methionine as well as N-terminal protein acetylation were defined as variable modifications. Raw files from the phosphoproteome and acetylome were processed together with the corresponding FP data in a separate parameter group with phosphorylation of serine, threonine, and tyrosine or acetylation of lysine, respectively, as additional variable modifications. Reporter ion MS3 was set as quantification type and TMT as isobaric labels.

Proteomics data analysis

For data analysis and visualization, mostly Microsoft Excel, GraphPad Prism 5, RStudio (version [4.0.2]), and the Perseus software (Cox and Mann, 2012; Tyanova et al., 2016) (v. 1.6.2.3) were used. Firstly, all reversed hits and proteins that were only identified by site were removed from the proteingroups.txt output. Then, the reporter intensities were log2 transformed and median-centered to the overall median of the respective dataset. Further, samples were adjusted with ComBat (Johnson et al., 2007) from the "sva" package (version [3.30.1]) to remove batch effects between replicates. Principal component analysis was performed on the intersection of the replicates with the 'factoextra' package (version [1.0.7]) and plotted with the 'ggplot2' package (version [3.3.2]). One-way ANOVA test with multiple testing correction (Benjamini-Hochberg, BH) was used to determine significance. A protein was classified as differentially expressed with an FDR<0.05 and a FC > 2 at one or more time points. For clustering, all differentially expressed proteins were row-wise z-scored and hierarchical clustered with the Perseus software. The KEGG enrichment analysis was performed with the 'clusterProfiler' (Yu and He, 2016; Yu et al., 2012) package (version [3.16.0]) with an FDR threshold set to 0.05. To derive TF-target relationships, differentially expressed proteins from each time point (relative to iPSC) were processed with TRRUST (version 2) (Han et al., 2018).

Phosphoproteomics and acetyl-IP data analysis

Reverse hits were removed from the Maxquant output and filtered for class I (localisation probability>0.75) P-sites or Ac-sites, respectively. The reporter intensities were log2 transformed and median-centered using the correction factors from the corresponding FP data. Batch effect correction was performed with ComBat and differentially expressed P-sites or Ac-sites, were determined like for the FP dataset (ANOVA: FDR<0.05 and FC > 2). The differentially expressed P-sites were further used for kinase-substrate relationship prediction with the networkin (Horn et al., 2014) web-tool.

scRNA-seq data analysis

For transcriptomics analyses, previously published scRNA-seq data (Camp et al., 2017) was used, in which the same 2D differentiation protocol with TkDA3-4 cells was performed. In that study, a total of 425 single-cell transcriptomes from multiple time points (iPSC = 80, DE = 70, HE = 113, IH = 81, and MH = 81 single cells) during hepatocyte differentiation were acquired. The normalized log2(FPKM) expression of all single cells at one time point were averaged to make it comparable with the bulk proteomics dataset.

QUANTIFICATION AND STATISTICAL ANALYSIS

The differentiation experiments were repeated independently two or three times, respectively, as outlined in the Method details. One-way analysis of variance (ANOVA) was calculated with the 'stats' package (version [4.0.2]) using the the aov() and TukeyHSD() functions in Rstudio (version [4.0.2]). Benjamini-Hochberg was applied to correct for multiple testing errors. A protein, a P-site, or an Ac-site was classified as significant with an FDR < 0.05 and a FC > 2. Statistical details can also be found in the according figure legends.