LEIBNIZ-INSTITUT
FÜR ARBEITSFORSCHUNG
AN DER TU DORTMUND

fakultät für chemie
und chemische biologie

Functional improvement of stem cell derived hepatocyte-like cells after targeted FXR gene regulatory network manipulation

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES DES DOKTORS DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
FAKULTÄT FÜR CHEMIE UND CHEMISCHE BIOLOGIE DER TECHNISCHEN UNIVERSITÄT DORTMUND

VORGELEGT VON

DAVID CHRISTOPH FEUERBORN, M. Sc.

GEBOREN IN DORTMUND, DEUTSCHLAND

DORTMUND 2022

1. GUTACHTER: PROF. DR. JAN G. HENGSTLER
2. GUTACHTER: PROF. DR. JÖRG RAHNENFÜHRER

Eidesstattliche Versicherung (Affidavit)

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Table of Contents

Summary

Primary human hepatocytes (PHH) are important for clinical therapy as well as for studies in pharmacology and toxicology. Hepatocyte-like cells (HLC) derived from pluripotent stem cells offer the perspective of an unlimited supply of PHH, however, genome- and proteome-wide analyses demonstrated that HLC still show major differences compared to PHH. More recently it was shown that, HLC express hepatocyte- and non-hepatocyte-associated genes within the same cells, indicating that HLC reside in a hybrid state that can be targeted by bioinformatics-guided intervention [1]. In this context, it remains to be clarified whether cell line or differentiation protocol-specific differences lead to comparable hybrid states and if the reported hybrid state is a common feature among HLC. In this thesis, HLC obtained by two different protocols from three different induced pluripotent stem cell lines (iPSC) were compared using genome-wide transcriptomics. Furthermore, it was demonstrated that interventions to improve HLC differentiation by targeting the FXR gene regulatory network (GRN) increased the expression of hepatocyte-associated genes and suppressed undesired non-liver genes in HLC, thereby increasing their similarity to PHH. However, this has yet only been shown on the transcriptomic level. Here, functional assays of bile acid secretion and lipid droplet formation were performed to confirm that an FXR targeting intervention strategy does indeed increase the similarity of HLC to PHH.

Zusammenfassung

Primäre humane Hepatozyten (PHH) sind ein wichtiges Instrument in der klinischen Therapie, der Pharmakologie als auch der Toxikologie. Hepatozyten-ähnliche Zellen (HLC) abgeleitet von pluripotenten Stammzellen emöglichen einen prinzipiell unbegrenzten Zugang zu PHH, jedoch zeigen Genom- und Proteom-weite Analysen, dass HLC noch immer große Unterschiede verglichen mit PHH aufweisen. Zudem zeigen neue Studien, dass HLC Hepatozyten- und nicht-Hepatozyten-assoziierte Gene innerhalb einer Zelle exprimieren [1]. In diesem Kontext muss weiterhin geklärt warden ob Zelllinien- oder Differenzierungsprotokoll-spezifsiche Unterschiede zu vergleichbaren Hybridzuständen führen und ob die bereits bekannten Hybridzustände eine gemeinsame Eigenschaft von HLC sind. In dieser Studie wurden HLC mit zwei verschiedenen Protokollen und drei verschiedenen induzierten pluripotenten Stammzelllinien (iPSC) generiert und mit Hilfe Genom-weiter Transkriptomanalysen verglichen. Des Weiteren wurde demonstriert, dass Interventionen zur Verbesserung von HLC Differenzierung die das FXR Genreuglationsnetzwerk zum Ziel haben die Expression von Hepatozyten-assoziierten Genen und die Unterdrückung von unerwünschten nicht-Lebergenen in HLC fördern können, wodurch die Ähnlichkeit mit PHH auf der Transkriptomebene steigt. Um zu zeigen, dass diese Veränderungen in der Genexpression, die durch die FXR Intervention gefördert wurden, zu nachhaltigen Verbesserungen der HLC führen, wurden in dieser Arbeit funktionelle Aspekte der Zellen in Bezug zur Gallensäurensekretion und Lipidtropfenformierung untersucht.

Abbreviations

%	Percent
°C	Degree Celsius
µg	Microgram
µL	Microliter
µM	Micromolar
µm	micrometer
ABCB	ATP-binding cassette
AdSC	Adult stem cells
AFP	α-Feto-protein
ALB	Albumin
ATP	Adenosine triphosphate
BMP	Bone morphogenic protein
BSEP	Bile salt export pump
CACO	Colorectal adenocarcinoma
CAR	Constitutive androstane receptor
CDCA	Chenodeoxycholic acid
CDX	Caudal type homeobox
ChiPSC	Cellartis human iPSC
CLD	Chronic liver disease
CNS	Central nervous system
CO_2	Carbon dioxide
Ct	Cycle threshold
CXCR	Chemokine receptor
CYP	Cytochrome P450
DBD	DNA binding domain
DE	Definitive Endoderm
DiPa	Differentiation Pattern
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid

ESC	Embryonic stem cells
FGF	Fibroblast growth factor
FOXA	Forkhead Box Protein
FXR	Farnesoid X Receptor
FXRi	FXR intervention
g	g-Force
GSK	Glycose synthase kinase
HGF	Hepatocyte growth factor
HHEX	Hematopoietic Homeobox Protein
hiPSC	Human iPSC
HLC	Hepatocyte-like cell(s)
HNF	Hepatocyte nuclear factor
HSC	Hematopoietic stem cells
HSC	Hepatic stellate cell(s)
iPSC	Induced pluripotent stem cell(s)
LBD	Ligand binding domain
LD	Lipid droplet
LDLR	Low-density lipoprotein receptor
LSEC	Liver sinusoidal endothelial cell
LXR	Liver X Receptor
mg	Milligram
MOI	Multiplicity of infection
mRNA	Messenger RNA
NANOG	Homeobox protein NANOG
ng	Nanogram
NPC	Non-parenchymal cell
NR	Nuclear receptor
NTCP	Na^+-taurocholate-cotransporting-PP
OCA	Obeticholic acid
OCT4	Octamer-binding transcription factor

OSM	Oncostatin M
PCR	Polymerase Chain Reaction
PDX1	Pancreatic And Duodenal Homeobox 1
PHH	Primary human hepatocyte(s)
POU5F1	POU domain class 5 transcription factor 1
PSC	Pluripotent stem cell(s)
PXR	Pregnane X Receptor
qRT-PCR	Quantitative real-time PCR
RNA	Ribonucleic acid
RT	Room temperature
RXR	Retinoid X receptor
SHP	Small heterodimeric partner
SOX	SRY-box
SOX	SRY-related HMG-box gene
STM	Septum transversum mesenchyme
TCA	Taurocholic acid
TF	Transcription factor
TG	Triglyceride
TGF	Transforming growth factor
WNT	Wingless Int-1
x	fold

List of Figures

1 Introduction

1.1 Stem cells and early human development

Stem cells can be defined as cells that preserve a capacity to self-renew and an ability to generate differentiated progeny [2]. Two central aspects of stem cell biology are potency and self-renewal [2]. Potency defines how many differentiated cell types a stem cell can produce and self-renewal is the capacity to proliferate in the same state [3]. Pluripotent stem cells (PSC) are capable to differentiate in cells of all three germ layers and correspond *in vivo* to embryonic stem cells (ESC) of the inner cell mass in the blastocyst stage of human development around days five to fourteen after conception [4]. Excluding extraembryonic tissues, PSC have the potential to form the entirety of cells of the human body and thus the study of induction, maintenance and exit of the self-renewing, pluripotent cell state is a focal point of developmental biology [3,4]. It has long been established that pluripotency is maintained by a certain set of transcription factors, which prevent differentiation and promote proliferation [5–7]. A key regulator of pluripotency is the transcription factor OCT4, which however does not act as a simple binary switch [8]. Rather, only medium expression levels of OCT4 in an interplay with tightly regulated expression of partner transcription factors keep stem cells locked in a pluripotent state [6]. In contrast, loss of expression as well as strong upregulation of OCT4 is connected with loss of pluripotency and commitment to other cell fates [9,10]. It is known, that transcription factor SOX2 stabilizes OCT4 expression and together, both proteins directly influence the expression of NANOG transcription factor, in turn creating a circuit that is essential for, but not sufficient to maintain the undifferentiated and self-renewing state of stem cells [8,11,12]. Thus, a vast body of basic science established an interconnected group of transcription factors which maintain pluripotency in stem cells, culminating in the subsequent discovery that ectopic expression of certain combinations of these factors is sufficient to reprogram somatic cells into stem cells [13]. In adult humans, PSC do not exist, rather tissue-specific multipotent stem cells termed adult stem cells (AdSC) that produce only a narrow set of differentiated progeny reside in certain niches [14]. The stem cell niche is an area of a tissue that provides a defined microenvironment in which other cells interact with stem cells either to maintain them or to promote their differentiation [14].

In these organs, AdSC have been found to constantly replenish the tissue with their progeny due to regular cell turnover or in case of tissue injury. A well-studied example of adult stem cells are hematopoietic stem cells (HSC) which reside in the bone marrow and can differentiate to all major cell types of the blood [15]. Further AdSC populations have been found in a number of, but not all organs, including brain, skin, intestine, heart and liver.

1.2 The naïve gut tube and early hepatogenesis

In triploblastic organisms, the gastrula is segregated into three germ layers: ectoderm, endoderm and mesoderm [16]. In a process termed neurulation the outer layer, the ectoderm, will differentiate to form the neural tube, neural crest cells and an epidermal region [16]. From these cells arises i.A. the central nervous system (CNS), the peripheral and enteric nervous system, the epidermis, nails, hair, the olfactory and oral epithelium and the eyes. Increasing evidence suggests that the middle layer, the mesoderm, and the innermost layer, the endoderm arise from a common progenitor cell population termed mesendoderm. Commitment to either mesoderm or endoderm lineage is achieved by a difference in Nodal signaling, where higher Nodal signaling promotes endoderm specification. The mesoderm continues to form the muscles, bones and connective tissue, the urinary as well as the reproductive and circulatory system. By a series of grafting experiments in mice it was shown that shortly after gastrulation anterior and posterior endoderm are not yet determined and when grafted posteriorly, anterior endoderm acquires posterior character and vice versa, showing regional identity to be dependent on signaling from the surrounding mesoderm [17,18]. This reciprocal signaling between definitive endoderm and mesoderm is essential for both cell populations and contributes to the development and patterning of an endodermal naïve gut tube, comprising several parenchymal progenitor cells of trachea, lung, esophagus, stomach, liver, pancreas and small intestine (**Figure 1.1**, left) [19] . In detail, morphogen gradients of WNT, BMP, FGF and retinoic acid which are secreted by the adjacent mesoderm in a temporally and spatially defined order, establish the regional identity of the endoderm by inducing expression of several transcription factors.

By that the gut tube can be organized into foregut, midgut and hindgut depending on expression of HHEX, PDX1 and CDX2 transcription factors respectively. Focusing on the foregut, labeling experiments have determined two bilateral populations of lateral endoderm and a smaller cell population at the ventral midline endodermal lip (VMEL) to be precursors of liver progenitors [20]. Early liver bud morphogenesis starts, when the cuboidal foregut epithelium which arose from these populations thickens to a columnar hepatoblast epithelium, which subsequently shifts to a pseudostratified epithelium [21].The hepatoblasts then start to delaminate and invade the adjacent septum transversum mesenchyme (STM), in turn forming the liver bud [22] (**Figure 1.1**, right).

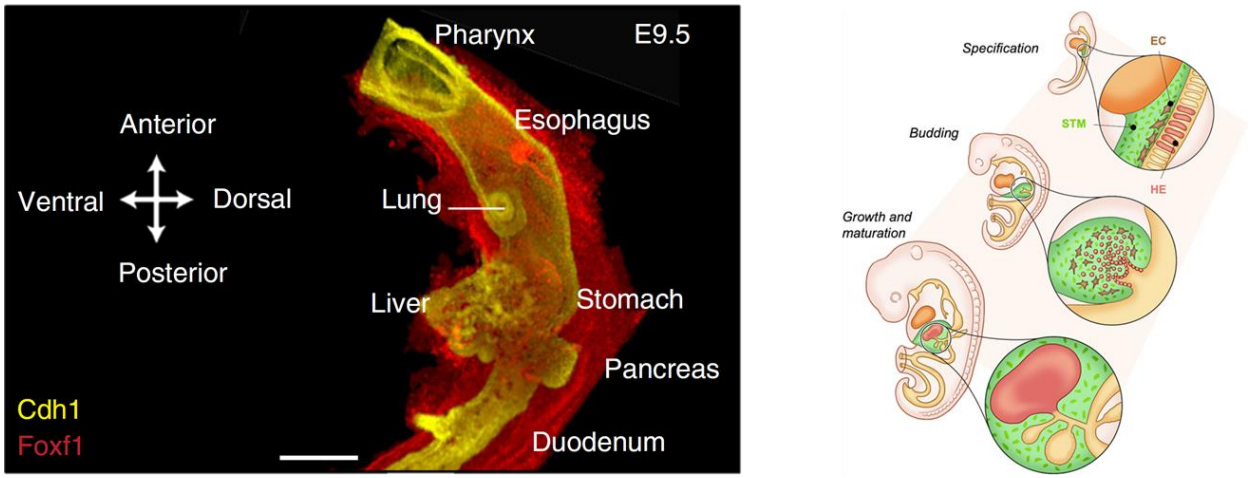


Figure 1.1: Development of the naïve gut tube and the liver bud. Whole mount immunostaining of mouse foregut (E9.5) with CDH1-positive endoderm (yellow) and FOXF1-positive surrounding mesoderm in red (left) [19]. Formation of the liver bud from specification and budding to growth and maturation (right) [23].

Adding to the complexity of liver development, soon after hepatoblasts delaminate into the STM, the fetal liver is invaded by hematopoietic progenitor cells and thus temporarily develops into the primary hematopoietic organ. Both, blood stem cells and developing liver cells have been shown to influence each other in this time frame. For instance, it was shown that oncostatin M (OSM), a cytokine produced by hematopoietic cells is substantial for successful differentiation of hepatic progenitor cells [24,25]. Liver hematopoiesis peaks in the second trimester of human development and then gradually declines until birth, concomitant with the migration of hematopoietic stem cells to the bone marrow, the principal residence in adult humans [26].

1.3 Structure, composition and function of the adult liver

The adult human liver is located in the right upper quadrant of the abdominal cavity, right next to the stomach, below the diaphragm and on top of the gallbladder [27]. As the largest gland in the human body, it is also the heaviest organ with a complex architecture and a central role in metabolic homeostasis, in which it is responsible for the metabolizing, synthesizing, storing and redistributing of nutrients, fats, carbohydrates and vitamins [28]. In addition, the liver produces major quantities of serum proteins, including ALB, as well as enzymes, cofactors and acute phase proteins [28]. Roughly 80 % of cells in the liver are parenchymal cells termed hepatocytes, and approximately 20 % are non-parenchymal cells (NPC), which consist of liver sinusoidal endothelial cells (LSEC), resident macrophages termed Kupffer cells, lymphocytes, leukocytes, hepatic stellate cells and cholangiocytes [29]. Blood flow to the liver is achieved by two major blood vessels: the hepatic artery and the portal vein that drain into the capillary system of the liver, the sinusoids [27]. The hepatic artery carries oxygenated blood from the aorta to a number of organs and accounts for roughly a quarter of hepatic blood flow [27]. However, the main blood flow to the liver is achieved by the portal vein, which drains oxygen-poor but nutrient-rich blood from the capillary system of the gastrointestinal tract into the sinusoids [28]. Blood arrives in the liver at a structure termed portal triad, where capillary branches of the hepatic artery, the portal vein and the bile duct are held together by connective tissue. Each portal triad is at the vertex of a hexagonal shape that comprises the functional unit of the liver, the lobule. The blood then drains to the central vein, which is situated in the center of the lobule and from which cords of hepatocytes radiate to the rim (**Figure 1.2**) [29]. Depending on their position along this axis, hepatocytes are exposed to different levels of oxygen, where oxygen concentration drops toward the central vein. In consequence, energetically demanding oxidative processes, such as gluconeogenesis or beta-oxidation, are performed by hepatocytes in oxygen-rich, periportal areas, whereas energetically less demanding processes like glycolysis, bile acid synthesis and xenobiotic metabolism are conducted in pericentral hepatocytes. This phenomenon is known as liver zonation and is only established in the first few weeks after birth [30].

Although it has been proven that oxygen saturation plays a crucial role in liver zonation, studies in mice showed that only a minor fraction of zonated genes is regulated by oxygen-sensitive transcription factors, whereas approximately a quarter of genes is regulated by WNT signaling [31]. Thus, before entering the blood circulation, nutrients and potentially harmful substances must pass the liver. Consequentially, the liver has adapted strong compensatory mechanisms to deal with damage, most prominently a unique capacity to regenerate [32,33].

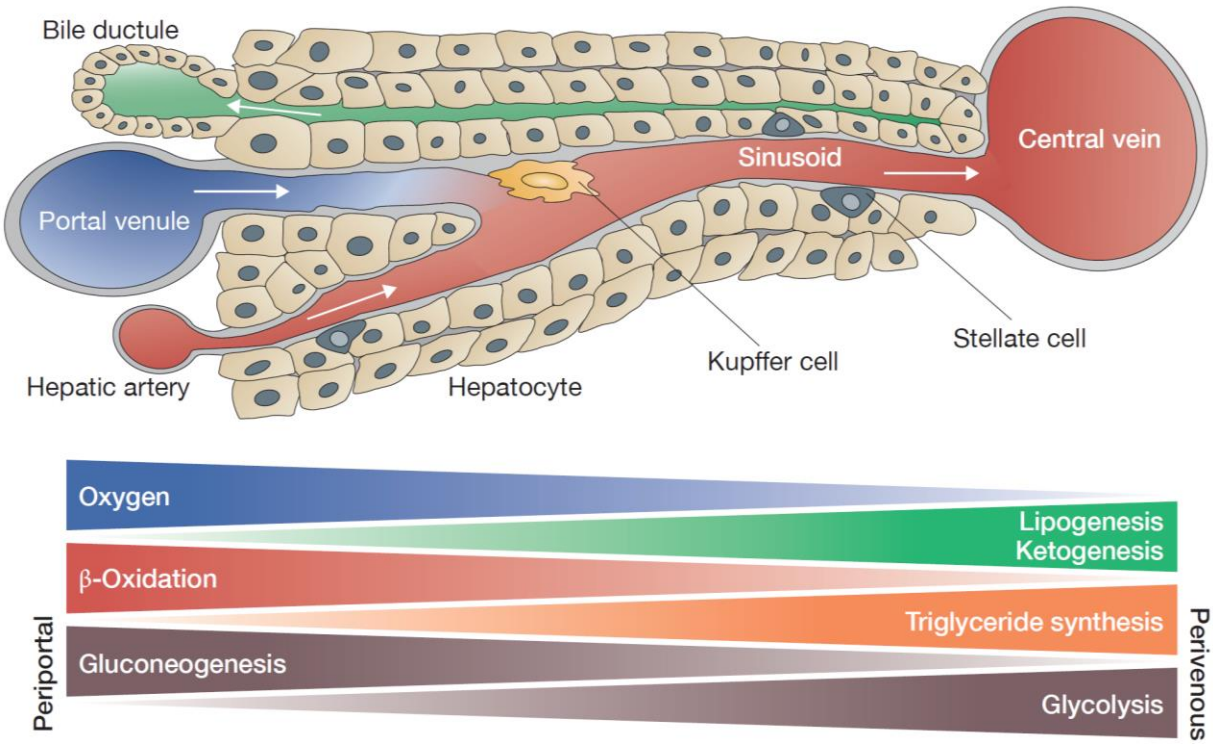


Figure 1.2: Structure and zonation of the porto-central axis of the liver lobule. Oxygen levels and metabolic processes differ in zones of the liver lobule. Energy demanding processes such as β-oxidation and gluconeogenesis occur in oxygen-rich periportal areas, whereas lipogenesis and glycolysis occur mainly in pericentral areas [34]**.**

In opposite direction to the blood flow, bile, a mixture primarily consisting of water (98 %) supplemented with bile acids, fats and inorganic salts, leaves the liver through the bile canaliculi, which drain into a ductular system that merges to form the common hepatic duct [29]. Together with the cystic duct from the gallbladder they form the common bile duct which joins the pancreatic duct and enters the duodenum [27]. Arriving in the duodenum the amphipathic bile salts assist lipase enzymes to digest lipids taken up by the organism from food by forming emulsifying micelles.

Approximately, 95 % of the bile acids that enter the duodenum are again absorbed by enterocytes and transported back to the liver, a process termed enterohepatic circulation [35].

1.4 Regulation of bile acid and lipid metabolism in hepatocytes

The diverse set of liver functions is associated with the tight interplay of numerous cell types assembled in a complex organ architecture that allows a structural-functional organization. Nevertheless, the major cell type of the liver are the hepatocytes, which control core functions including the maintenance of molecules like glucose/glycogen, bile acids, triglycerigdes, cholesterol and vitamins A and D. Furthermore, hepatocytes metabolize heme and bilirubin, regulate pH by urea synthesis and provide clotting factors. An important prerequisite for the correct execution of many hepatocyte functions is a high degree of cell polarization. Uniquely, the apical domain of one hepatocyte forms a tubular structure with a neighboring hepatocyte and thus contributes to a network of canaliculi, which serve as smallest branches of the biliary tree [29]. In contrast, the basolateral domain of hepatocytes consists of the sinusoidal membrane, where hepatocytes are in contact with the blood plasma, and the lateral plasma membrane, which builds connections with adjacent hepatocytes by desmosomes and gap junctions. Given the unequal tasks performed over the plasma membranes on the canalicular and on the sinusoidal sides, both membranes exhibit major differences in composition and organization. The sinusoidal membranes for instance are dominated by high-affinity cell surface receptors sensing growth factors, hormones and cytokines from the blood. Moreover, it is rich in bile acid uptake transporters like the sodium taurocholate cotransporter (NTCP) and lipid and iron scavenging receptors such as the low-density lipoprotein receptor (LDLR). In contrast, the canalicular plasma membrane is enriched in bile acid efflux transporters, such as the bile salt export pump (BSEP), involved in ATP-dependent transport of bile acids to the canaliculi [36].

Therefore, a complex routing process of resident membrane proteins to canalicular and sinusoidal membranes is vital for hepatocyte function and necessitates an extensive ER and Golgi network, which is also needed for the numerous proteins produced in hepatocytes and destined for secretion. Regulation of gene expression in hepatocytes is predominantly accomplished by a class of ligand dependent transcription factors, termed nuclear receptors (NR) [29]. Usually, upon binding of a ligand to the ligand binding domain (LBD) NR tend to homo- or heterodimerize and, if they are not already expressed in the cell core, translocate to the nucleus, where they bind in promoter regions using their DNA binding domain (DBD) [29]. One nuclear receptor, the farnesoid X receptor (FXR) is activated by certain bile acid metabolites such as chenodeoxycholic acid (CDCA) or taurocholic acid (TCA) as well as synthetics and semi-synthetics like GW4064 or obeticholic acid (OCA) [35]. Upon activation, FXR translocates to the nucleus where it dimerizes with cooperating transcription factors such as retinoid X receptor (RXR) and binds to responsive elements on the DNA in promoter regions of genes causing up- or downregulation. Primarily, FXR binds to genes associated with lipid and bile acid homeostasis in liver, including the small heterodimer partner (SHP) and fibroblast growth factor 19 (FGF19) [37]. Both of which repress enzymes linked in the classical pathway of bile acid synthesis from cholesterol like cytochrome P450 7A1 (CYP7A1) and cytochrome P450 8B1 (CYP8B1). Therefore, accumulation of bile acids in hepatocytes leads to a negative feedback mechanism starting with FXR by repressing expression of bile acid synthesis related cytochromes. In addition, FXR directly upregulates bile salt transporters, such as the bile salt export pump (BSEP) involved in clearance of bile salts from the cytoplasm to the canaliculi. FXR is also involved in regulation of autophagic processes in hepatocytes [38]. Autophagy is a degradative process that keeps up intracellular homeostasis under conditions of oxidative stress, starvation or nutrient-rich environments. As it degrades lipid droplets (LD), autophagy also influences lipid metabolism of the cell by promoting lipolysis. Several recent studies have shown, that FXR suppresses autophagy in nutrient-rich, high fat environments and thus causes the upregulation of triglyceride (TG) contents in hepatocytes [38]. Therefore, dependent on intracellular state the nuclear receptor FXR influences a various set of functions in hepatocytes, including repression of bile acid synthesis and suppression of lipophagy.

1.5 Stem cell technology in pharmaceutical research and regenerative medicine

Despite its remarkable regeneration capabilities, the liver is still susceptible to a number of diseases. As a progressive deterioration of liver functions chronic liver disease (CLD) is a major burden for healthcare systems worldwide, causing approximately two million deaths per year [44]. Risk factors include alcohol consumption, obesity, diabetes and drug abuse, but also viral hepatitis can progress from acute liver damage to chronic liver disease. CLD is a continuous process involving inflammation, damage and regeneration of hepatocytes leading to fibrosis and eventually end-stage liver disease known as cirrhosis. The final stage of the disease is associated with loss of liver architecture, appearance of regenerative nodules, a vascular reorganization and a deposition of the ECM and is considered irreversible [45]. At a cellular level fibrosis and subsequent cirrhosis are connected to activation of hepatic stellate cells (HSC), which represent a main source of ECM during liver injury [46]. If the organ can no longer perform its physiological functions, the patient experiences liver failure and needs a liver transplant. However, donor livers are in short supply and although successful transplantation is the only known remedy for liver failure it is linked to a loss of life quality and life expectancy. In consequence, there is not only an enormous demand for transplantable hepatic material due to the current rise in CLD, but also an increase in the demand for alternative treatments for diseases leading to CLD. However, a major flaw in the drug development process are pre-clinical experiments in isolated *in vitro* systems of cells or tissues and animal models, which are not predictive of drug efficacy in humans [39]. Unsurprisingly, roughly 90 % of newly developed drugs fail in late stage clinical trials due to their lack of efficacy in their intended disease indication [39]. Thus, there is a high demand for technologies, including stem cell-based systems, which allow to increase predictability by enabling accurate disease-modeling in pre-clinical studies. Yet, all published organoid liver systems for pharmacological as well as regenerative purposes suffer from a set of shortcomings common to most of the currently published artificial organ systems, including off-target gene activation or repression, absent or reduced vascularization, low reproducibility and scaling issues [40]. Additionally, the liver combines a complex architecture with a multitude of essential functions unparalleled in other organs of endoderm origin thus creating a substantial challenge.

1.6 Current pluripotent stem cell technologies

In the past, access to pluripotent stem cells was limited by ethical considerations concerning isolation of embryonic stem cells and their derived stem cell lines [41]. The discovery, that a certain set of transcription factors when expressed in somatic cells could restore pluripotency, opened up the possibility of an unrestrained availability of PSC [13]. These induced pluripotent stem cells (iPSC) have since become a key technology that allows an unlimited, ethically acceptable supply of stem cells that have been used alongside ESC lines to obtain differentiated progeny of the three primary germ layers *in vitro* to gain insights into developmental processes or to utilize them in clinical applications [42–44]. In principle, ESC as well as iPSC can be generated from every human and have therefore the potential to be used in a personalized medical approach, enabling autologous cell therapy and accurate disease modelling [45,46]. Although there have been major advances in differentiation protocols for distinct cell types, the recreation of the stage dependent and dynamically changing microenvironment of the developing human *in vitro* has become a major challenge in stem cell biology. In consequence, current iPSC-derived cells are often immature or suffer from off-target differentiation and therefore lack functional capabilities in comparison to their primary counterparts [47–49]. In contrast to the modelling of other *in vivo* processes, *in vitro* differentiation is hampered by the fact that cells from important developmental stages in human fetal organ development are not legally available and thus cannot serve as reference in directed differentiation approaches. Therefore progress in the field has been largely informed by studies conducted in animal models [19]. Nevertheless, when available, human fetal cells have profoundly impacted understanding of organogenesis *in vivo* and *in vitro*, especially when used in a highly time resolved manner [47,50,51]. Even though the possibilities of iPSC for medical applications seem wide, a current massive limitation of iPSC-derived cells in cell therapy is their tendency to retain some of the stem cell like properties acquired by reprogramming [42]. Early on, it was discovered that the expression of the reprogramming factor OCT4 is a feature that PSC share with some cancer cells [52]. Unsurprisingly, cells overexpressing OCT4 and cooperating transcription factors, transplanted into organisms often develop neoplasms if expression is not tightly regulated [53]. Additionally, it was shown in mice that overexpression of OCT4 during

reprogramming is associated with off-target gene activation and epigenetic aberrations, indicating an imperfect reprogramming process, which is at least partly responsible for poor differentiation outcome compared to ESC-derived cells [54]. Thus, some skepticism about the applicability of current iPSC for regenerative purposes remains.

1.7 Pathways to pluripotent stem cell derived hepatocyte-like cells

Currently much effort is directed towards the derivation of hepatocyte-like cells, which represent a possible replacement for the use of primary human hepatocytes in clinical therapy, pharmacology and toxicology [1,55,56]. Through the study of early development in model organisms it is known that hepatoblasts, progenitors of hepatocytes, spring from cell populations of the lateral foregut endoderm. Consequently, the first step in *in vitro* approaches to differentiate towards hepatocyte cell fate is the differentiation of PSC into a near-homogenous population of endoderm cells [42,47,57]. Studies in model organisms showed that members of the transforming growth factor beta (TGFβ) superfamily such as Activin A and Nodal of morphogens are key to achieve endoderm specification [43]. Complexes of downstream effectors SMAD2/3 then recruit activating and inhibiting transcription co-factors leading to cell-type specific responses [43]. Additionally, WNT agonistic signals, either by WNT3A or synthetic substances such as GSK3β-inhibiting CHIR99021, help cells to achieve the competence to respond to Activin A signaling. Cell populations derived by several variations of the ActivinA/WNT3A approach have succeeded to express genes that are representative of definitive endoderm including CXCR4, FOXA2 and SOX17 [1,47]. Furthermore, these cells have been used, with a varying degree of success, for in vitro differentiation of cells from the endoderm lineage comprising pancreas, intestinal and liver cells [58,59]. To mimic signals coming from the developing heart and mesenchyme in a second step towards hepatocyte cell fate, researchers often employ combinations of BMP4 and FGF2 [60]. In parallel, downregulation of pluripotency genes and increase in hepatic competence of endoderm cells has also been achieved by DMSO treatment [61]. In agreement with in vivo studies, resulting cell populations show expression of marker proteins AFP, HNF1A and HNF4A similar to hepatoblasts [42,43,60]. In a final step of differentiation, a majority of

approaches uses combinations of recombinant HGF and OSM growth factors resembling paracrine signaling in the developing liver for hepatic maturation [42,47]. Initially, OSM is secreted by blood stem cells during the hematopoietic phase in liver development [62]. Aided by dexamethasone, OSM binds to the OSM receptor, prompting a more hepatocyte-like morphology and inducing functional maturation and suppression of fetal markers like AFP [62]. In contrast, HGF has been proven beneficial in hepatic differentiation because of its association with adult liver processes of liver regeneration, hepatocyte growth and maturation [63]. A schematic overview of the most common *in vitro* differentiation approach is given below (**Figure 1.3**).

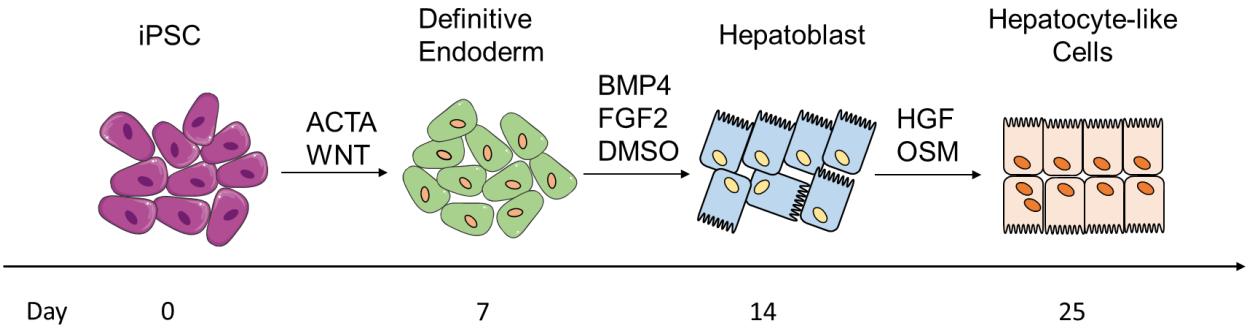


Figure 1.3: In vitro differentiation of iPSC into hepatocyte-like cells. Differentiation of stem cells into HLC is achieved via initial induction of definitive endoderm by Activin A and activation of WNT signaling followed by combination of BMP2/FGF2 or alternatively low DMSO concentrations to initiate the hepatic progenitor phase. The length of the protocol and the time intervals of the major phases may vary between different protocols but usually is set to 18 to 25 total differentiation days.

1.8 Current hepatocyte-like cells suffer from a transcriptional hybrid state

Despite remarkable progress in in vitro differentiation of stem cells, the capabilities of e.g. hepatocyte-like cells have in the past been routinely overestimated. Independent of the differentiation protocol of origin, HLC usually show a lack of expression in gene clusters connected to basic hepatocyte functions such as xenobiotic metabolism [42]. Major transcription factors of the nuclear receptor class (PXR, FXR, CAR, and LXR) regulating essential enzyme expression (BSEP, CYP) are generally underrepresented, resulting in cells that are functionally immature compared to their primary counterparts [1]. Besides underrepresented factors, HLC also express genes, such as the fetal albumin analog AFP, that are rather related to hepatic progenitor cells, highlighting the level of immaturity of these cells [47]. Even more importantly, besides gene expression that is associated with hepatic development, HLC additionally increase expression of genes that are characteristic of other lineages and are not related to hepatic development [42,64]. Interestingly, misguided differentiation is not only limited to endodermal lineages but also includes expression of genes found in cellular offspring derived from ectoderm and mesoderm [1]. In the light of these findings the capacity of HLC in their current state to be used as a substitute for PHH in toxicological studies as well as in cell therapies is questionable. To investigate the transcriptional state of HLC various studies have employed sequencing approaches in combination with algorithms that are able to compare the entirety of gene regulatory networks in HLC with that of PHH and reference liver tissue [64,65]. Here, HLC regularly achieve at maximum 75 % of the network level of PHH, yet additionally also 50 % of diverging lineages, such as intestine [42]. The high number of genes from undesired lineages naturally raised the question whether HLC in vitro exist in subpopulations or exhibit a mixed identity that is characterized by transcriptional and functional overlap of two or more cell lineages within one cell [1]. Here, evidence currently accumulates that HLC indeed exhibit mixed identities within one cell and efforts are made to understand this possible hybrid state in engineered cell types [66]. Although, a wealth of publications has centered on improving HLC by various interventions, usually involving one or more critical transcription factors, most studies have failed to determine the transcriptional and functional impact of their intervention strategies on the hybrid state [48,49].

Supported by single cell RNA sequencing a recent study first highlighted the potent hybrid state of current HLC and explored possible routes to reduce the misguided differentiation outcome [1]. This lead to a series of proof-of-principle experiments that demonstrated how overexpression and activation of a single transcription factor sufficed to not only increase the hepatic transcriptional state, but also suppressed undesired intestinal state attributes and thus lead to a reduced hybrid state of HLC in general [1] (**Figure 1.4**). Notably, until now, there is no evidence supporting the hybrid state to be a feature of fetal hepatocytes *ex vivo* [1].

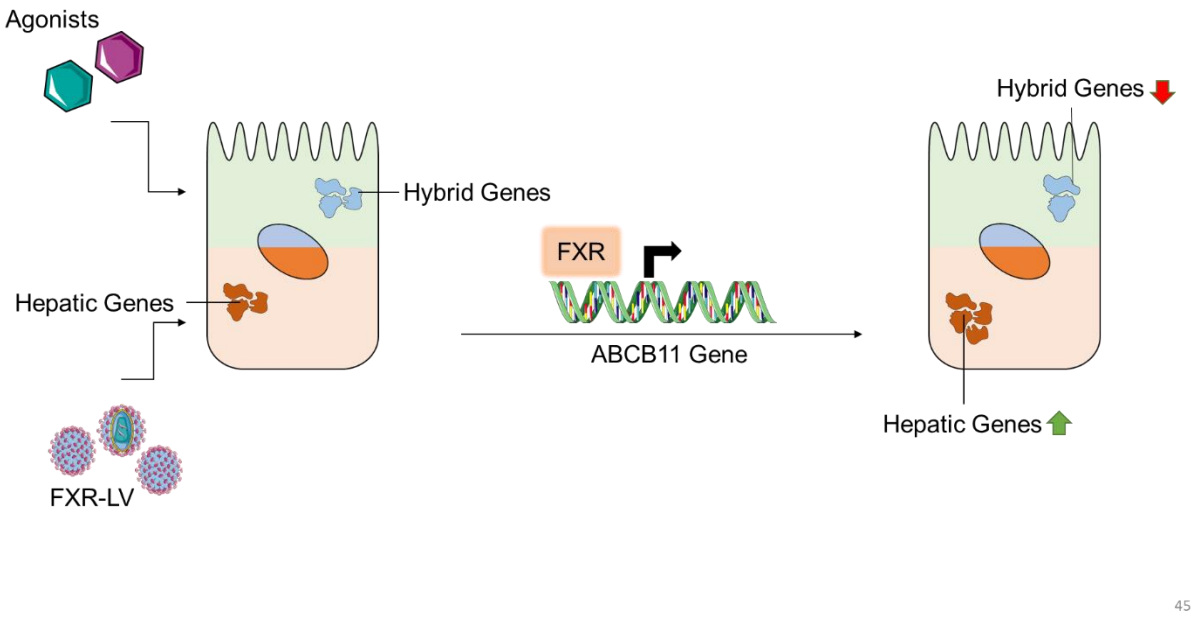


Figure 1.4: Influence of FXR intervention on HLC. Overexpression and activation of nuclear receptor FXR leads to a robust upregulation of hepatic genes and a mild downregulation of undesired hybrid genes.

1.9 Aim of the thesis

An increasing amount of evidence suggests that hepatocyte-like cells derived from pluripotent stem cells either are immature in comparison to their *in vivo* counterparts or suffer from an aberrant gene expression that entraps them in a hybrid cell state. A recent publication proposed not only that these cells exhibit a potent cellular hybrid state between a hepatic and an intestinal cell fate, but also that the nuclear receptor FXR plays a role in pushing HLC towards a hepatocyte cell state and concomitantly reduces hybrid state cell character [1]. Although convincing sequencing data suggests an improvement, it remained unclear to which degree these interventions have enhanced FXR-related cellular functionality of HLC and, given reported problems of reproducibility, whether these results can be reliably repeated in stem cells from other donors and in HLC derived in other protocols.

2 Materials and Methods

2.1 Materials

2.1.1 Equipment

Table 1: Equipment used in this study.

Item	**Manufacturer**
5075 ELV Autoclave	Tuttenauer
AF 100 ice machine	Scotsman
7500 Fast & 7500 Real-Time PCR System	Applied Biosystems
Agilent 2100 Bioanalyzer	Agilent Technologies
Balance EW	Kern
Biofuge Heraeus Fresco 21	ThermoScientific
Centrifuge 5424R	Eppendorf
Centrifuge MEGA STAR 1.6R	VWR
Chemical Safety cabinet Herasafe™	Heraeus
CO_2 Incubator C150 R Hinge 230	Binder
Diaphragm Vacuum Pump	Vacuumbrand
Dual-Action Shaker KL2	Edmund Bühler
Freezer (-20°C)	Siemens
Freezer (-80°C)	ThermoFisher Scientific
HiSeq2500 Sequencer	Illumina
IKAMAG RCT magnetic stirrer	IKA
Laminar Flow Hood	Waldner
LSM 880 confocal laser scanning microscope	Zeiss
Megafuge 1.0R	Heraeus
MiniSpin / Minispin Plus	Eppendorf
NanoDrop 2000	PeQLab Biotechnologie
Neubauer Counting Chamber	VWR
PH-meter CG 842	Schott
Pipette Reference	Eppendorf
Pipette Research Plus	Eppendorf

Item	Manufacturer
Pipetteboy	Integra
Precision Balance EW 150-3M	Kern
Rocking Platform Shaker	VWR
Shaker KS 260 basic	IKA
SONOPULS mini20 sonicator	Bandelin
Testtube Shaker	VWR
Vortex Genie 2	Bender&Hobein
Waterbath GFL 1083	Ges. f. Labortechnik

2.1.2 Consumables

Table 2: Consumables used in this study.

Item	Company	ID
Biosphere Filtered Tip 100 µL	Sarstedt	70.760.212
Biosphere Filtered Tip 1000 µL	Sarstedt	70.762.211
Biosphere Filtered Tip 20 µL	Sarstedt	70.1116.210
Biosphere Filtered Tip 200 µL	Sarstedt	70.760.211
Cell Scraper	Sarstedt	83.183
Falcon tube 15 mL	Sarstedt	62.554.512
Falcon tube 15 mL	Sarstedt	62.547.254
Filtropur v100 1000 mL Vacuum Filter	Sarstedt	83.1824.001
Filtropur v100 250 mL Vacuum Filter	Sarstedt	83.1822.001
Filtropur v100 500 mL Vacuum Filter	Sarstedt	83.1823.001
Kimtech Science Delicate Task Wipes	Kimberly Clark	7216
MicroAmp Optical 96-well Reaction Plate	Applied Biosystems	N801-0560
MicroAmp Optical Adhesion Film	Applied Biosystems	4311971
PCR SingleCap 8er-SoftStrips 0.2 mL	Biozym Scientific	710988
Pipette Tips 1000 µL	Sarstedt	70.762
Pipette Tips 200 µL	Sarstedt	70.1116
Pipette Tips 100 µL	Sarstedt	70.760.002

Item	Company	ID
RNase-free Microfuge Tubes 1.5 mL	ThermoFisher Scientific	AM12400
RNaseZap® RNase Solution	ThermoFisher Scientific	AM9780/AM9782
SafeSeal 0.5 mL microtube	Sarstedt	72.699
SafeSeal 1.5 mL microtube	Sarstedt	72.706
SafeSeal 2.0 mL microtube	Sarstedt	72.695.500
Serological Pipette 5 mL	Sarstedt	86.1253.001
Serological Pipette 10 mL	Sarstedt	86.1254.001
Serological Pipette 25 mL	Sarstedt	86.1685.001
Serological Pipette 50 mL	Sarstedt	86.1689.001
Tissue Culture Flas 25 cm^2	Sarstedt	83.3910.502
Tissue Culture Flas 75 cm^2	Sarstedt	83.3911.502
Tissue Culture Plate Flat Bottom 12-well	Sarstedt	83.3921
Tissue Culture Plate Flat-Bottom 24-Well	Sarstedt	83.3922
Tissue Culture Plate Flat-Bottom 6-Well	Sarstedt	83.1839

2.1.3 Reagents

Table 3: Reagents used in this study.

Item	Company	ID
(2'Z,3'E)-6-Bromoindirubin-3′-oxime	Sigma-Aldrich	B1686-5MG
Acetic acid	Carl Roth	3738.5
Alexa Fluor 488 Donkey anti-Mouse	Dianova	715-545-150
Alexa Fluor 488 Donkey anti-Rabbit	Dianova	715-545-150
Amino Acid Solution	PAN Biotech	SO-33100
Ampure XP beads	Beckman Coulter	A63881
Anti-FXR	NovusBio	NB400-153
Anti-BSEP	Invitrogen	PA5-27742
β-Mercaptoethanol	ThermoFisher Scientific	31350010

Item	Company	ID
BSA	Carl Roth	8076.4
BSA fatty acid free	Sigma-Aldrich	
$CaCl_2$	Sigma-Aldrich	5239.1
DEF-CS 500 Culture System	Takara Bio Europe	Y30010
Definitive Endoderm Differentiation Kit	Takara Bio Europe	Y30035
Hepatocyte Differentiation Kit	Takara Bio Europe	Y30050
Cell Tracker Green CMFDA Dye	ThermoFisher Scientific	C2925
Chenodeoxcholic Acid	Sigma-Aldrich	C9377
Chloroform	Carl Roth	7331.2
DAPI	Invitrogen	D3571
DEPC Treated Water	Invitrogen	750024
Dexamethasone	Sigma-Aldrich	D4902-25MG
Direct-zol RNA MiniPrep Plus kit	Zymo Research	R2071
DMSO	Sigma-Aldrich	34869
DPBS with Ca^{2+} and Mg^{2+}	ThermoFisher Scientific	14040133
DPBS without Ca^{2+} and Mg^{2+}	ThermoFisher Scientific	14190144
Ethanol	Merck	100983
Gentamycin	PAN Biotech	P06-13001
GlutaMAX Supplement	ThermoFisher Scientific	35050061
GW4064	Sigma-Aldrich	G5172
High Capacity cDNA Reverse Transcription Kit	Applied Biosystems	436
Insulin-transferrin solution (ITS)	Sigma Aldrich	I3146
Knockout DMEM	Life Technologies	10829-018
Knockout Serum Replacement	ThermoFisher Scientific	10828010
L-Glutamine	Pan Biotech	P04-82100
Laminin 111	Biolamina	LN111-02
Laminin 521	Biolamina	LN521-25

Item	Company	ID
Lyophilized Rat-tail Collagen	Roche	11179179
MEM NEAA	ThermoFisher Scientific	11140050
MinElute PCR Purification Kit	Qiagen	28004
Nextera DNA Library Prep Kit	Illumina	20025519
Nextera XT Library Prep Kit	Illumina	FC-131-10
Penicillin/Streptomycin	PAN Biotech	P06-07100
Phalloidine-Rhodamine	ThermoFisher Scientific	R415
Potassium Chloride	Carl Roth	6781.1
Potassium dihydrogen phosphate	Carl Roth	3904.1
Oleic Acid	Sigma-Aldrich	O1383
Qiazol Lysis Reagent	Qiagen	79306
ROCK Inhibitor	Sigma-Aldrich	Y0503-1MG
Roti Histofix 4 %	Carl Roth	P087.5
Sera Plus	PAN Biotech	3702-P103
Sodium Chloride	Carl Roth	3957.2
TaqMan Universal Master Mix	Applied Biosystems	4440038
Transferrin Human	Sigma-Aldrich	T8158-100
Triton X-100	Carl Roth	3051
TRIzol	ThermoFisher Scientific	15596026
Trypan Blue	Sigma-Aldrich	T6146
TrypLE	ThermoFisher Scientific	12563011
Tween 20	Sigma-Aldrich	P7949
Tween 80	Sigma-Aldrich	P8074

2.1.4 Cells

Item	Company	ID
ChiPSC18 induced pluripotent stem cells (male)	Takara Bio Europe	Y00305
ChiPSC22 induced pluripotent stem cells (male)	Takara Bio Europe	Y00325
JHU106i	WiCell	P106
Primary Human Hepatocytes (male)	BioReclamationIVT	M00995-P
CACO-2 cell line	ATCC	HTB37

2.2 Methods

The following method section of my thesis contains parts which I contributed to the publication in the Journal of Hepatology under "*Identification of an FXR-modulated liver-intestine hybrid state in iPSC-derived hepatocyte-like cells*" [1].

2.2.1 Preparation and maintenance of induced pluripotent stem cell cultures

Induced pluripotent stem cells were purchased from Cellartis (ChiPSC18, ChiPSC22, JHU106, Takara Bio Europe) and stored in liquid nitrogen at -150°C. Culture vessels were prepared with COAT-1 solution (Cellartis) diluted 1:20 in ice-cold DPBS plus Ca^{2+} and Mg^{2+}, using 0.1 mL to 0.15 mL per cm^2 of the culture vessel. The culture vessels were incubated for at least 30 min at 37°C. Cells were thawed in a water bath at 37°C and transferred to DEF-CS full medium (Cellartis) containing DEF-CS growth factor 1(diluted 1:333), 2 (diluted 1:1000), 3 (diluted 1:1000) reagents (Cellartis) Subsequent to centrifugation at 300 x g for 5 min, supernatant was aspirated and the cell pellet resuspended in 5 mL of full DEF-CS full medium. Immediately before the transfer of the cell suspension to the prepared culture vessels, coating solution was aspirated. Cells were left for attachment at 37°C and after 24 h medium was changed to DEF-CS medium with DEF-CS growth factors 1 (diluted 1:333) and 2 (diluted 1:1000) (DEF-CS maintenance medium). From here on daily medium changes were conducted with DEF-CS maintenance medium until full confluency and stem cell compaction. To keep the stem cell cultures clear of unwanted spontaneous differentiation, cultures were regularly checked for morphological signs of differentiation. Additionally, immunostaining against as well as qRT-PCR for pluripotency markers such as POU5F1 and NANOG were routinely conducted. For splitting and passaging of cells, confluent and compacted cultures were washed once with DPBS(-/-) containing no Mg^{2+} and Ca^{2+}. Cell dissociation was achieved by treating iPSC with 0.015 mL per cm^2 TrypLE select (ThermoFisher) at 37°C for 5 min. To obtain a single cell suspension, 10 volumes of DEF-CS full medium was added and the cell suspension collected. Using a Neubauer hemocytometer cells

2.2.2 Differentiation of hiPSC to definitive endoderm

Culture vessels to be used for the derivation of definitive endoderm from three hiPSC cell lines (ChiPSC18, ChiPSC22, JHU106) cells were coated 60 min at RT with 0.1 mL per cm² Definitive Endoderm Differentiation Coating Reagent (Cellartis). Fully confluent hiPSC cells were dissociated, counted and collected. After centrifugation (300xg, 5 min) and resuspension in Definitive Endoderm D0 medium, differentiation toward definitive endoderm was started by seeding 34000 cells/cm² in the newly coated culture vessels. After 24 h medium was changed Definitive Endoderm D1 medium. This was continued according to the manual schedule until day 7 of differentiation. Further details on the Cellartis® Definitive Endoderm Differentiation system, Cat.no. Y30035, can be found in the manufacturer's instructions.

For differentiation to definitive endoderm according to the Hay Protocol cells were treated as previously described with only minor changes. [1] Briefly, culture vessels were coated with a 1:20 dilution in PBS of Laminin 521 and Laminin 111 (Biolamina) in a 1:3 ratio. Cell culture vessels were incubated with Laminin or Cellartis hepatocyte differentiation coating solutions for 60 min at 37°C. Upon reaching full confluence and showing a dense monolayer, ChiPSC18 cells were harvested and counted. Next, the cell suspension was aliquoted as required for seeding 50,000 cells/cm2 and centrifuged at 300xg for 5 min, the supernatant was aspirated, the pellet was dissolved in DEF-CS medium containing growth factors 1,2 and 3 reagents and transferred to the cell culture vessel. Definitive endoderm differentiation was induced on the next day by addition of Hay definitive endoderm priming medium containing Activin A (100 ng/mL, Peprotech) and WNT3A (50 ng/mL, R&D). The endoderm priming medium was renewed every 24 h. On day 3 of the differentiation the endoderm priming medium was changed to endoderm priming medium with only Activin A (100 ng/mL).

2.2.3 Differentiation of hiPSC-derived definitive endoderm to hepatocyte-like cells

To start differentiation of definitive endoderm to hepatocyte-like cells, culture flasks were flooded with 0.1 mL per cm^2 of a 3:1 ratio mix of Laminin 111 and Laminin 521 solution (both Biolamina), followed by a 1:20 dilution in DPBS. Efficient coating of the cell culture vessels was achieved in an incubator at 37°C for 60 min. Definitive endoderm cells were washed with DPBS(-/-) containing no Mg^{2+} and Ca^{2+}, incubated with 0.1 mL / cm^2 TrypLE Select at 37°C for 3 min. By adding one volume of DPBS with 10 % FBS the cell suspension was collected. Cells were counted using a Neubauer hemocytometer and centrifuged (300xg, 5 min) before resuspension in Thawing and Seeding medium at $2.6x10^5$ cells / mL. Cells were seeded with a density of $1.3x10^5$ cells / cm^2 and incubated at 37°C for 48 h. Medium changes until day 25 of differentiation were conducted according to the manufacturer's instructions. At day 25 HLC were harvested for downstream applications

For differentiation of hiPSC-derived definitive endoderm to hepatocyte-like cells according to the Hay protocol cells were treated as previously described [67]. Briefly, culture of cells differentiated to definitive endoderm by the Hay Protocol (see above) was continued with KSR/DMSO differentiation medium on day 6 of the differentiation for 5 days. For the first three days medium was changed daily and then on the fifth day of differentiation. On day 11 of the differentiation, medium was switched to HepatoZYME maturation medium containing HGF (10 ng/mL, Peprotech) and OSM (20 ng/mL, Peprotech). Medium was changed every 48 h. Upon completion of day 25, HLC were harvested for downstream applications.

2.2.4 Lentiviral transduction and agonist treatment

The lentiviral vector construct pLV[Exp]-EGFP:T2A:Puro-EF1A>hNR1H4 [NM_001206993.1] carrying ORFs for EGFP and FXR as well as the resulting lentiviral particles were generated in collaboration with VectorBuilder (Chicago, USA, vector map in the appendix, **Supplementary figure 5.1**) as previously described [68]. Quantitative PCR of a fragment of the ENV region of the vector amplified from genomic DNA of transduced HEK293 cells showed the viral titer to be $2.19x10^9$ transducing units / mL. HLC differentiated according to the CEL protocol were transduced on day 13 of differentiation while HLC differentiated according to the HAY protocol were transduced on day 10 of differentiation. For transduction, medium was aspirated and replaced with a third of the volume of a usual medium change. Then, the HLC were exposed to lentiviral particles at a multiplicity of infection (MOI) of 10 and a polybrene concentration of 5 µg/mL and incubated for 16 h at 37°C. After incubation, medium was switched to the next medium according to schedule. The efficiency of the lentiviral transduction was assessed by EGFP fluorescence. For the treatment of HLC with agonists, DMSO stock solutions of CDCA (100 mM, Sigma) and GW4064 (1.5 mM, Sigma) were prepared. Cells were exposed to 100 µM of CDCA and 1.5 µM of GW4064 with the routine medium change on day 22 and again on day 24 of differentiation. At day 25 of differentiation HLC were either fixed and used in immunostaining or total RNA was collected.

2.2.5 Cultivation of primary human hepatocytes.

Cryopreserved primary human hepatocytes (PHH) were purchased from BioIVT. Four different donors were used for data generation of RNA sequencing, for qRT-PCR and for immunostaining. For cultivation of PHH instructions of a published standard operating procedure were followed with minor modifications [69]. For the coating, lyophilized rat-tail collagen (Roche Diagnostics) was dissolved in 40 mL 0.2 % acetic acid at 4°C overnight. Standard 12-well plates were shortly exposed to 1 mL collagen solution (250 µg/mL). Then, plates were left to dry overnight under sterile conditions and before use were washed at least three times with PBS to remove residual acidity. In a water bath at 37°C, PHH were thawed and immediately transferred to a tube with full culture medium consisting of 1 % Penicillin/Streptomycin, 10 µg/mL gentamycin, 10 ng/mL ITS, 2 mM L-glutamine, 100 nM dexamethasone and 10% FCS in Williams E medium.

Cells were counted using a Neubauer hemocytometer and approximately 65000 cells/cm^2 were seeded and left for attachment at 37°C for at least 3 h. During the first 10 minutes of attachment the cultures were gently shaken to enable equal distribution. After attachment, PHH were washed three to four times with PBS, the attachment medium was replaced with culture medium without 10 % FCS and cells were incubated at 37°C until further use.

2.2.6 Cultivation of CACO-2 cells

A stock of cryopreserved CACO-2 cells were purchased from ATCC. Thawing was performed at a water bath at 37°C, resuspended in 5 mL DMEM and finally centrifuged at 300xg for 5 min. The supernatant was aspirated, the cell pellet dissolved in 10 mL fresh DMEM and the cells counted using a Neubauer hemocytometer. In a cell culture vessel 50000 cells/cm^2 were seeded and afterwards incubated at 37°C. Medium was switched to fresh DMEM twice a week and cells passaged upon achieving approximately 80 % confluence. Before downstream usage of the cells, CACO-2 was allowed to reach full confluence while avoiding overgrowth.

2.2.7 Total RNA isolation and cDNA synthesis for q-RT-PCR

For total RNA collection, cells were either detached from the culture vessel by TrypLE with subsequent centrifugation or directly treated with Qiazol reagent (Qiagen) and scraped off the surface of the vessel. Cells were collected in small tubes and sonicated at 50 % intensity for 30 s (5 s interval, 2 s break) with a sonicator (SONOPULS mini20, Bandelin). For storage RNA lysates were frozen at -80°C.

For total RNA isolation, samples taken up in Qiazol were defrosted on ice and treated 1/5 volume of chloroform. Samples were shaken gently for 15 s, incubated for 2 min at RT and centrifuged for 15 min at 12000 rpm and 4°C. Next, the aqueous phase including the RNA was taken up in 1/2 volume of isopropanol, the lower phase was discarded. After gentle shaking the samples were then incubated for 10 min at RT, centrifuged for 15 min at 12000 rpm and 4 °C. The resulting supernatant was discarded, the RNA pellet washed with 1 volume of ice-cold absolute EtOH. Samples were vortexed, centrifuged for 5 min at 12000 rpm and 4 °C and the cells subsequently washed with 1 volume of 75 % EtOH. Again, the samples were vortexed and centrifuged for 5 min at 12000 rpm and 4 °C. For air drying any residual liquid was discarded and samples were incubated for 5 min at RT with open lid. The pellets were resuspended in 1/10 volume of DEPC H_2O and shortly vortexed. The concentration of the samples were measured at the NanoDrop system. RNA isolates were stored at -80 °C.

RNA was isolated from hiPSC lines of ChiPSC18, ChiPSC22, JHU106 from the iPSC stage (Day 0 of differentiation), from the DE stage (Day 7 of differentiation) and from the HLC stage (Day 25 of differentiation). For FXR plus agonist intervention experiments RNA was isolated from iPSC: ChiPSC18 (CEL protocol, D0, n=3), ChiPSC18 (HAY protocol, D0, n=2), ChiPSC22 (CEL protocol, D0, n=2), JHU106 (CEL protocol, D0, n=3); definitive endoderm: ChiPSC18-DE (CEL protocol, Day 7, n=3), ChiPSC18-DE (Hay protocol, Day 5, n=2), ChiPSC22-DE (CEL protocol, D7, n=2), JHU106-DE (CEL protocol, D7, n=3). For control HLC: ChiPSC18-HLC (CEL protocol, D25, n=3), ChiPSC18-HLC (HAY protocol, D25, n=2), ChiPSC22-HLC (CEL protocol, D25, n=2), JHU106-HLC (CEL protocol, D25, n=3); FXR plus CDCA and GW4064 treated HLC: ChiPSC18-HLC (CEL protocol, D25, n=3), ChiPSC18-HLC (HAY protocol, D25, n=2), ChiPSC22-HLC (CEL protocol, D25, n=2), JHU106-HLC (CEL protocol, D25, n=3) and PHH (D0, n=4). For cDNA synthesis, 1 µg of isolated RNA was transcribed using a high-capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific) according to manufacturer's instructions.

2.2.8 Quantitative Real-Time PCR

Quantitative real-time PCR (qRT-PCR) was conducted on a 7500 Rea-Time PCR System (Applied Biosystems) using 25 ng of cDNA, TaqMan Universal Master Mix (ThermoFisher Scientific) and the following primer probes: ALB (Hs00609411_m1), HNF4A (Hs00604435_m1), HNF1A (Hs00167041_m1), NR1H4 (Hs01026590_m1), ABCB11 (Hs00994811_m1), POU5F1 (Hs04260367_gH), NANOG (Hs04399610_g1), CXCR4 (Hs00607978_s1), SOX17 (Hs00751752_s1), FOXA2 (Hs00232764_m1), CDX2 (Hs01078080_m1), KLF5 (Hs00156145_m1), ISX (Hs01368145_m1). Stepwise conditions for each PCR were: 50°C for 2 min, 95°C for 10 min, 40 cycles of 15 s at 95°C and 1 min at 60°C. For analysis, the $2^{-\Delta\Delta Ct}$ method was applied, where GAPDH was used as the housekeeping gene and the expression levels were nomalized in reference to the iPSC population at day 0 of the differentiation.

2.2.9 Immunofluorescence

Cells were fixed with 4 % PFA for 15 min at 37°C. Subsequently, cells were washed at least twice to get rid of residua PFA. On an orbital shaker, cells were washed, permeabilized using Triton X-100 for 10 min, again washed for 5 min and treated for 1 h with a blocking solution consisting of 10 % BSA and 0.1 % Tween-20 in PBS, followed by overnight incubation with antibody solution containing 3 % BSA, 01. % Tween-20 and the primary antibody at a 1:200 dilution in PBS. Primary antibodies used were anti-ALB (HPA031024, 1:200, Sigma), anti-AFP (MAB1369), anti-CDX2 (ab114247, 1:200, Abcam), anti-ISX (HPA060328, 1:200, Atlas Antibodies), anti-AGR2 (HPA007912, 1:200, Atlas Antibodies), anti-DPPIV (ab215711, 1:200, Abcam) and anti-HNF4A (HPA004712, Atlas Antibodies). On the following day, cells were washed 3x5 min on an orbital shaker and incubated with antibody solution consisting of 3 % BSA, 0.1 % Tween-20 and the secondary antibody at a 1:100 dilution in PBS. Subsequently, the cells were incubated for 2 h under light protection. Depending on the experiment, the secondary antibodies used were Alexa Fluor 488 Donkey anti-Rabbit (1:100), Alexa Fluor 488 Donkey anti-Mouse and ´Cy3-AffiniPure Donkey anti-Rabbit (all Jackson Immuno Research Laboratories). For staining of cell nuclei, the cells were incubated with a solution containing 1:10000 DAPI for 10 min at RT. For staining cytoskeletal actin cells were additionally incubated with a solution containg 1:5000 phalloidin-rhodamine (Life Technologies) for 10 min at RT. Before imaging, cells were washed 3x5 min with PBS and stored for maximally 4 days at 4°C.

2.2.10 Comparative analysis of CMFDA secretion into bile canaliculi

On day 25 of differentiation, HLC with and without FXRi were mounted on an LSM 880 confocal microscope (Zeiss) at 37°C and 5% CO_2. Per technical replicate, 3 fields of view were selected for control HLC and HLC with FXRi each and captured. Next, 5 µM of CMFDA was added to each well, starting with the control well and carefully avoiding delay. Immediately after addition, a time series capture was performed for the selected fields of view to ensure comparability based on the same imaging conditions. CMFDA fluorescence accumulation curves were measured by defining regions of interest (ROIs) over canaliculi-like structures in time-lapse movies and determining the change in mean intensity over time in the ROIs.

The temporal mean intensity profiles were then independently fitted to a single-exponential association equation derive the first-order time constant and amplitude of CMF fluorescence:

$$I_t = I_b + A\left(1 - e^{\left(-\frac{t-t_d}{\tau}\right)}\right)$$

Where:

I_t = Intensity at time t

I_b = background intensity offset

A = amplitude of CMFDA accumulation at saturation

t_d = delay before exponential association begins

τ = first order time constant

The time constants indicate the rate of CMFDA activation and export, while the amplitude indicates the amount of CMFDA that is activated and exported before the hepatocyte transport activity is saturated.

2.2.11 Oleic acid treatment and Adipored staining of HLC LD and image analysis

Oleic acid complexed with BSA (2 mol oleic acid / mol albumin, O3008, Sigma Aldrich) was added to fresh medium on day 25 of HLC differentiation to a concentration of 0.8 mM and provided to HLC with and without FXRi. For BSA control, medium was supplemented with 0.4 mM BSA. For the untreated control group, HLC with or without FXRi were provided with maintenance medium only. Cells were incubated for 48 h at 37°C and 5% CO_2. Next, cells were washed twice with PBS followed by addition of 10 µL of Adipored™ reagent (Lonza) per mL of PBS and 15 min of incubation at 37°C, 5% CO_2. Lipid droplet imaging was performed on a Leica SP8 confocal microscope equipped with a white light laser and an incubation chamber maintained at 37°C with 5% CO_2. Fluorescence excitation was set to 514 nm, with the emission detected between 536-672 nm on a HyD detector through a 25X 0.95 NA water immersion objective. Image resolution was set to 0.3 µm x 0.3 µm per pixel. Tile scanning and stitching was utilized to obtain accumulated fields of view of 700 µm x 700 µm per biological replicate. Cell segmentation was performed on the AdipoRed channel using the Cellpose machine-learning algorithm including the ‘cyto’ model with the following adjustments to parameters: [Estimated Diameter = 100 pixels, Flow threshold = 2.0, Cell Probability Threshold = -6.0] [70]. Lipid droplet segmentation was also performed on the AdipoRed channel using Cellpose, but with the ‘nuclei’ model and the following adjustments to the parameters: Estimated Diameter = 10 pixels, Flow threshold = 2.0, Cell Probability Threshold = -6.0]. Lipid droplet label images and Cell label images were then utilized to cross-reference and assign every detected lipid droplet to a cell. Lipid droplet parameters such as mean intensity and area were quantified through the regionprops function included in scikit-image [71]. All parameters were obtained as a data frame and used further for plotting and statistical analyses. Statistical analysis and visualization was performed in R, using a Wilcox rank-sum test for assessment of statistical significance of the non-parametric data.

2.2.12 mRNA-seq library preparation

Total RNA of cells in QIAzol reagent were collected as described before. To extract RNA the Direct-zol RNA MiniPrep Plus Kit (Zymo Research) including DNAseI treatment was applied. In a first step of library preparation a minimum of 100 ng total RNA were used for reverse transcription with 5 cycles of pre-amplification. Then, the Nextera DNA library prep kit (Illumina) was applied, in which 8 cycles of enrichment PCR were conducted, followed by a purification step with 0.9xAmpure XP beads.

2.2.13 Processing and basic analysis of RNA sequencing data

To remove 3`ends with base quality below 20 and adapter sequences, reads were trimmed by applying the Trim Galore package (v0.4.2). RNA sequencing reads were aligned to GRCh38 using STAR with the per sample 2-pass mapping strategy [72]. To detect PCR duplicates MarkDuplicate from Picard tools was employed (https://broadinstitute.github.io/picard/). Estimation of gene-wise read counts for RNA sequencing data were done by Gencode release 30 (GRCh38.p12) using RSEM [73].

Then, expression values were normalized as log CPM+1 (counts per million). Genes having an average log CPM < 0.1 were excluded. The 1000 most variable genes were used for Principal Component Analysis (PCA).. To identify differentially expressed genes, the DEseq algorithm was used to calculate scaling factors and estimate the dispersion [74].

2.2.14 Supervised clustering of differentially expressed genes

Differentially expressed genes were further analyzed by a supervised clustering approach, which relies on the definition of biologically interpretable differentiation pattern groups (DPG) that contain groups of genes displaying similar gene expression changes during iPSC to HLC differentiation in relation to the reference population, PHH. Representative sets of cutoffs defining DPGs were chosen after evaluating downstream analysis based on DPGs from several iterations of the procedure with a range of cutoff sets. For visualization, the log2 fold change of HLC over PHH was plotted against the log2 fold change of iPSC over PHH in a 2-dimensional space.

2.2.15 Downstream enrichment and further analyses of DPG

The differentiation pattern groups (DPG) yielded from the supervised clustering approach were further investigated by several downstream enrichment analyses. Tissue identity enrichment analysis of DPG was conducted with the help of the TissueEnrich tool v1.8 [75]. Furthermore, the function enrichGO from the package Clusterprofiler v3.16 was applied to investigate whether genes in certain DPG have an enrichment in certain gene ontologies [76]. The CellNet algorithm, which is based on the reconstruction of cell-type specific gene regulatory networks (GRN) [77]. The GRN status is assessed based on the weighted sum of z-scores of the genes, which belong to the gene regulatory networks. Calculated values of the GRN status range from 0 to 1, where higher values indicate that the activity of genes in the analyzed query sample is similar to the activity of genes in the respective reference tissue type. The network influence score (NIS) is part of the CellNet package and can be used to estimate the importance of transcriptional regulators for certain dysregulated GRN. Here, the weighted sum of z-scores of the genes of the networks of a tissue type is combined with another weighting of the transcription factor itself based on the number of target genes influenced by it.

Using Fisher tests, it was possible to quantify the statistical significance for comparisons of fractions of DEG in DPG. Finally, overlap ratios were generated. For the pair-wise overlap ratio score the function $OR=(O*N)/(n_1*n_2)$ is used, where O is the number of genes in the overlap, N represents the total number of genes and n_1 and n_2 are the number of differentially expressed genes under condition 1 and condition 2, respectively. Here, overlap ratio scores of 1.0 indicate random overlap, whereas values higher than 1.0 indicate an overlap that is higher than expected by chance in case of independence.

3 Results

The following results have already been published in the Journal of Hepatology under "*Identification of an FXR-modulated liver-intestine hybrid state in iPSC-derived hepatocyte-like cells*" [1].

3.1 Genome-wide analysis corroborates mixed cell identity in HLC

3.1.1 Differential gene expression analysis shows large similarity between different HLC

Previously it was shown, that HLC derived from iPSC of the type ChiPSC18 exhibit a hybrid state, in which a homogenous population of cells expresses not only markers of the intended hepatic differentiation, but additionally markers associated with unintended intestinal differentiation. However, given the range of protocols and cell lines applied to deriving HLC and reported problems of reproducibility, it was uncertain whether these results could be repeated over a wider range of HLC derived from various iPSC lines as well as HLC produced by other protocols. Thus, iPSC of the type ChiPSC18, ChiPSC22 and JHU106i were differentiated to HLC according to a proprietary protocol of Cellartis Takara Bio Europe. In addition, cells of the type ChiPSC18 were differentiated according to a published differentiation protocol [67].Next, an RNA sequencing series was performed, including samples of HLC derived from three different iPSC lines (ChiPSC18, ChiPSC22, JHU106), the corresponding iPSC and DE populations, three donor populations of PHH as well as three donor populations of epithelial colon tissue. With the resulting data (**Supplementary figure 1)**, differential gene expression analysis according to the Deseq2 algorithm was performed (**Supplementary figure 2)** in order to recreate a supervised clustering approach introduced by Nell et al, visualized in the differentiation pattern plot (**Figure 3.1**). Here, genes are plotted according to their expression fold changes of iPSC as the starting cell population and HLC as the end cell population in comparison to PHH as the target cell population. In this way, all genes are distributed in differentiation pattern groups (DPG 0-10). DPG 0 combined genes that remain constant over the course of the differentiation and have the same expression in HLC and PHH.

DPG 1 and DPG 6 on the diagonal have the same expression in iPSC and HLC and are thus not influenced by the differentiation protocol and remain at too low or too high expression levels, respectively.

Genes in DPG 2 and DPG 7 show a positive development during differentiation by approaching expression levels measured in PHH but continue to be insufficiently up- or downregulated. Favorable development is observed in DPG 3 and DPG 8, where genes are up- or downregulated close to levels of PHH. Genes that show an overall negative development over the course of differentiation are found in DPG 4 and 9 as well as DPG 5 and 10. Here, DPG 4 and 9 combine genes that exhibit excessive upregulation or downregulation. In addition, DPG 5 and DPG 10 include genes, whose expression should go down but goes up and vice versa. These clusters therefore represent misguided differentiation.

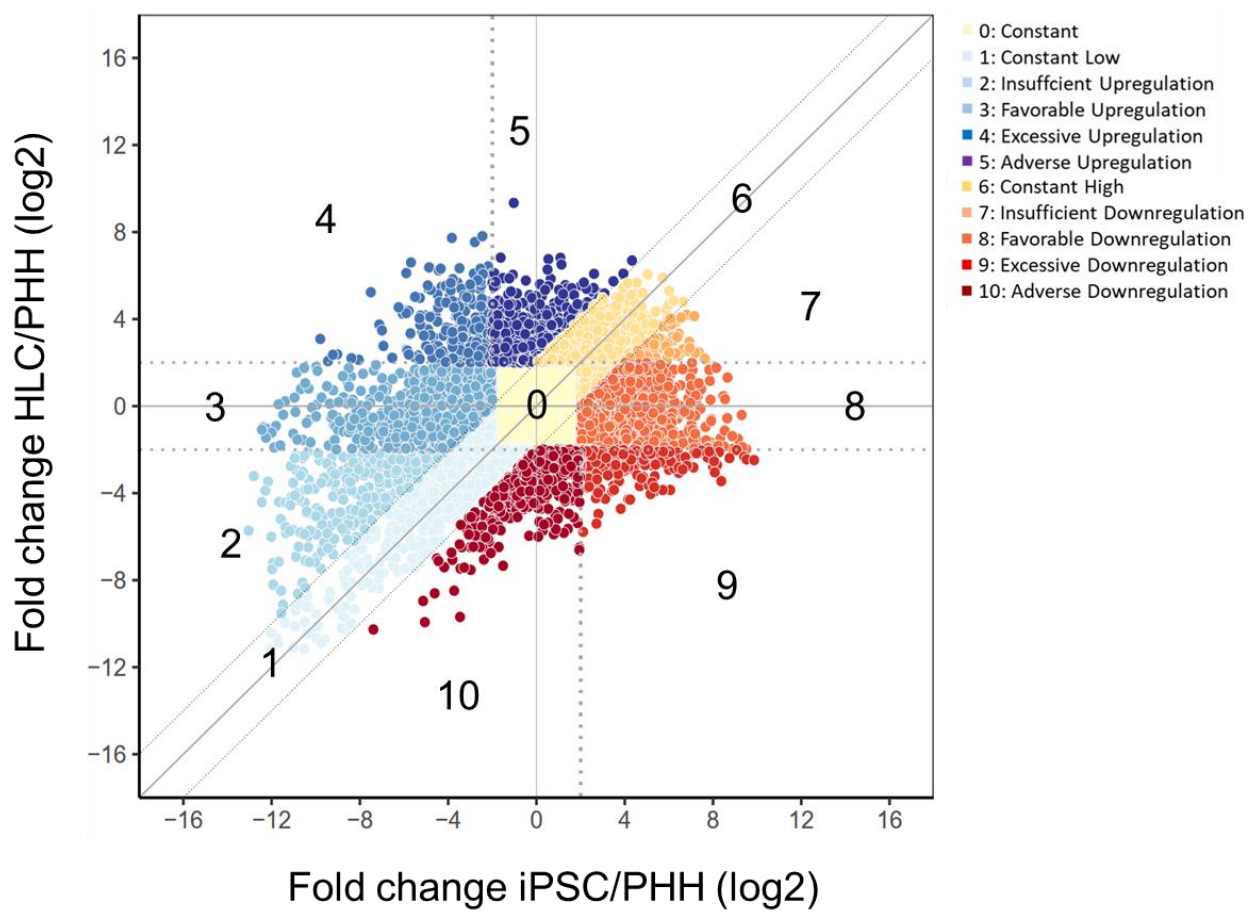


Figure 3.1: Differentiation pattern plot of JHU106 cell line. Supervised logical clustering of differentiation trajectories obtained from RNA-sequencing data, visualized in the differentiation pattern plot (DiPa plot). The x-axis represents log2 fold changes of iPSC over PHH, the y-axis indicates log2 fold changes of HLC over PHH. Dotted lines represent cutoffs of the clustering approach.

The differentiation pattern plot (DiPa plot) was reproduced for the ChiPSC18- and ChiPSC22-derived HLC and their corresponding iPSC populations (Supplementary Figure 5.1 and 5.2). To compare number and distribution of genes, DiPa plots were compared and showed largely the same distribution across all cell lines and the additional protocol. Furthermore, the number of genes in each DPG corresponded with the number of genes in DPG of other plots. Here, the vast majority of genes have the same expression in HLC and PHH and therefore remain in DPG 0 at the center of the plot. The remaining genes are distributed to DPG 1 to 10 (**Figure 3.2**).

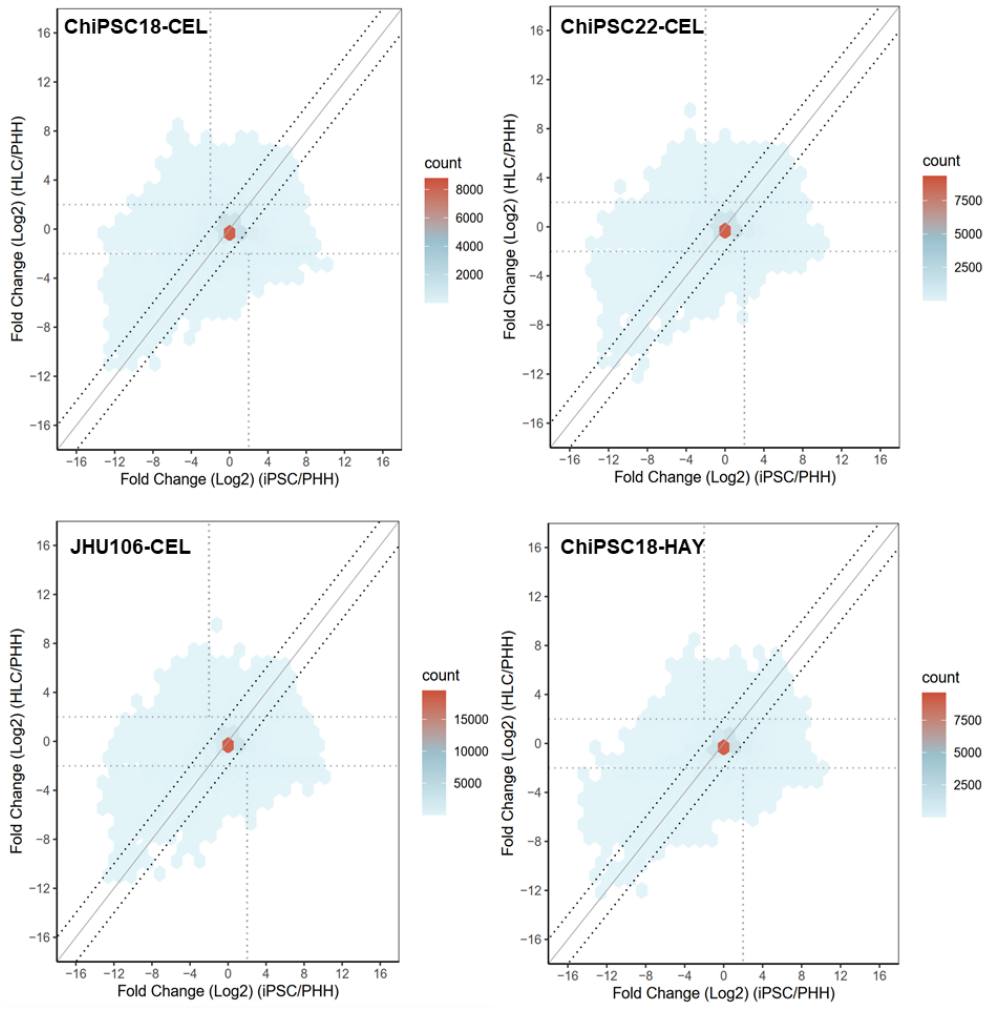


Figure 3.2: Differentiation pattern density plots. Differentiation pattern plots for each cell line (ChiPSC18, ChiPSC22, JHU106) and protocol (CEL, HAY) indicating the distribution of genes across all DPG.

3.1.2 Enrichment analysis demonstrates overlap across various HLC

Subsequently, several techniques were applied to characterize the genes in individual DPG across HLC of different origins. Using a tissue enrichment analysis tool, tissue group-enriched genes in DPG were investigated [78]. Analysis for insufficiently upregulated genes of DPG 1 and DPG 2 showed an overrepresentation of liver- and kidney-enriched genes. Furthermore, gene ontology enrichment analysis provided an insight into the processes most prevalent in DPG 2 [76]. Here, many metabolic processes were listed that are usually conducted in hepatocytes, such as small molecule metabolism, alpha amino acid metabolism or steroid metabolism. Importantly, enrichment results were comparable across different cell lines and protocols (**Figure 3.3**).

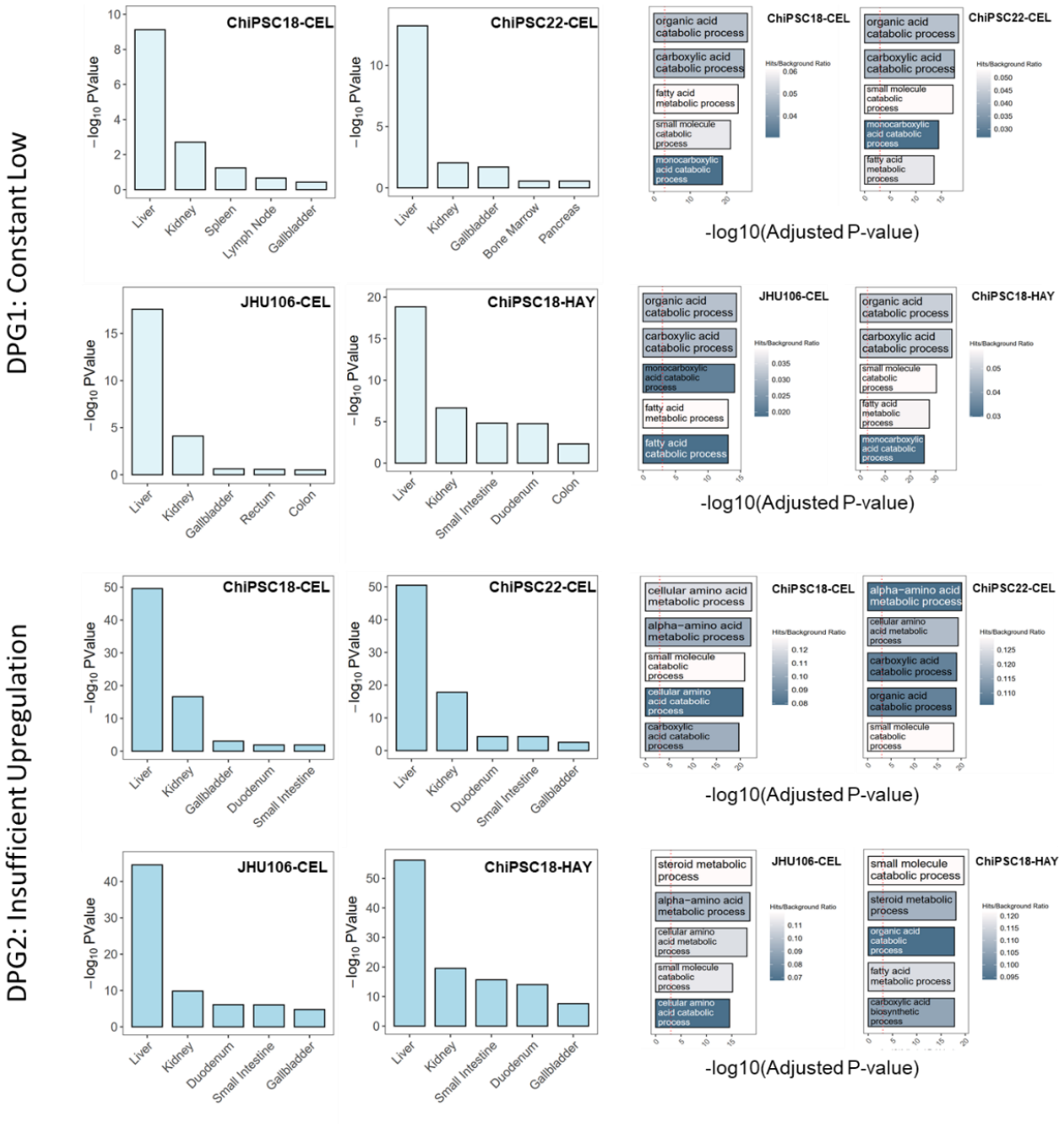


Figure 3.3: Tissue group and gene ontology enrichment analysis for DPG1/DPG2. Tissue group enrichment analysis (left) and gene ontology enrichment analysis (right) for DPG1 and DPG2 per cell line (ChiPSC18, ChiPSC22, JHU106) and protocol (CEL, HAY).

Investigating misguided differentiation, DPG 4 showed strong enrichment for intestinal genes across all HLC derived in the CEL protocol. Accordingly, these HLC also displayed a massive ontology enrichment for genes involved in digestion processes. For HLC derived in the HAY protocol a weaker but also stable tissue enrichment for intestinal lineages was found, yet gene ontology enrichment analysis revealed different processes such as response to TNF and LPS to be dominant. In DPG 5 cells differentiated according to the CEL protocol showed an enrichment in intestine-specific genes. For ChiPSC18-derived HLC the enrichment was weaker and independent of the differentiation protocol used. Here, CEL derived cells showed only poor enrichment for intestine and kidney, while cells differentiated according to the HAY protocol showed only poor enrichment for esophagus and skin. Expectedly, poor tissue enrichment among genes in DPG 5 of ChiPSC18-derived HLC was also reflected in gene ontologies that were again rather enriched for inflammatory processes such as response to TGF-ß stimulation. Taken together, analysis revealed a large overlap in enrichment for HLC derived in the CEL protocol. For HLC derived in the HAY protocol DPG of misguided differentiation almost always displayed diverging enrichment in tissue and gene ontologies, indicating that differentiation with these protocols achieve different states of differentiation and may therefore be misguided differently. However, the enrichment of intestinal identity among excessively upregulated genes (DPG 4) indicates that the acquisition of a hybrid state of mostly liver and intestinal identity is a shared feature between the CEL and HAY protocols. Interestingly, gene ontology enrichment of misguided clusters in HLC differentiated according to the HAY protocol often pointed towards inflammatory processes (**Figure 3.4**). Complete tissue enrichment and gene ontology enrichment tables for all DPG in all cell lines and both protocols can be found in the appendix under **supplementary figures 3-10**.

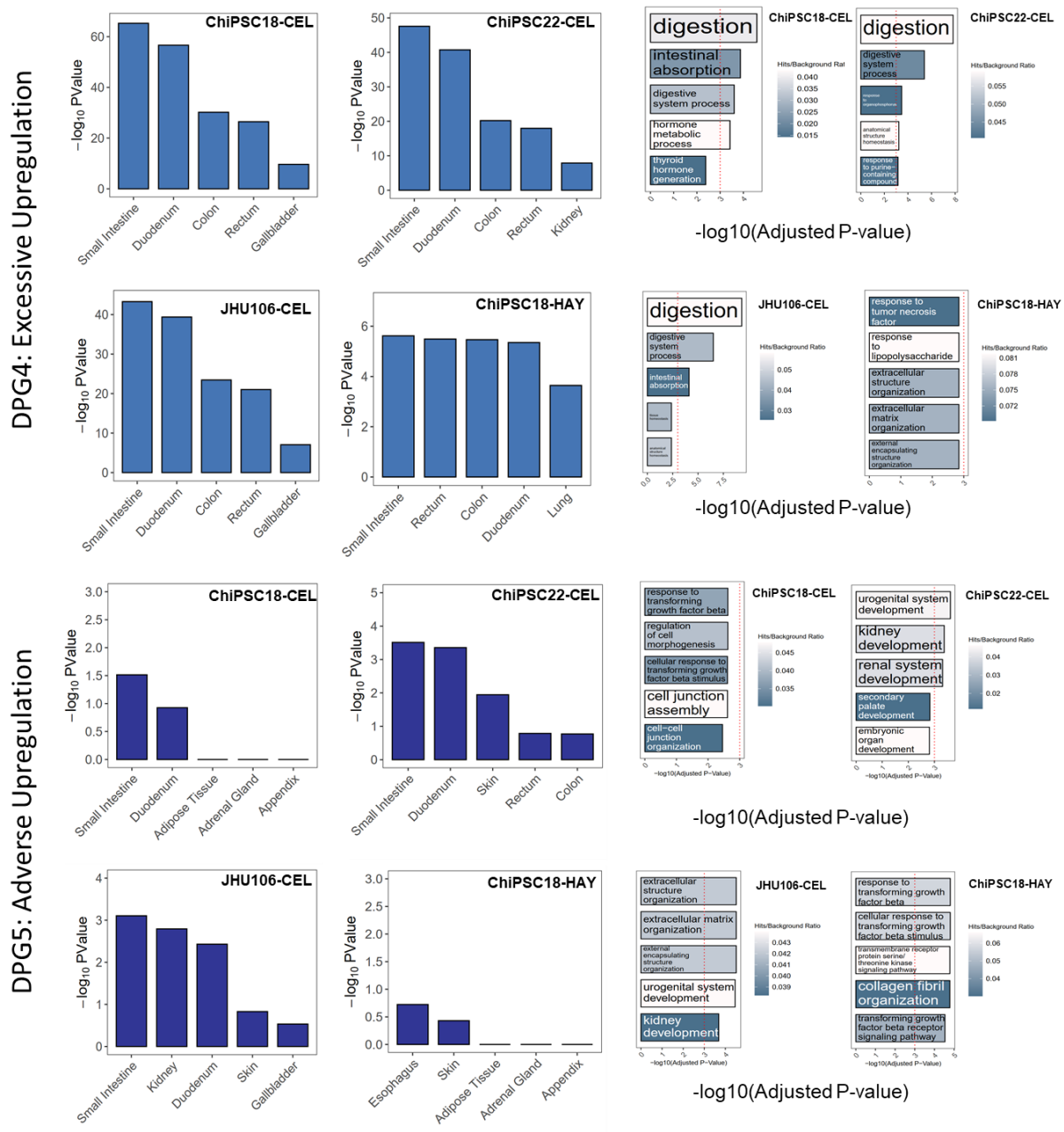


Figure 3.4: Tissue group and gene ontology enrichment analysis for DPG4/DPG5. Tissue group enrichment analysis (left) and gene ontology enrichment analysis (right) for DPG4 and DPG5 per cell line (ChiPSC18, ChiPSC22, JHU106) and protocol (CEL, HAY).

3.2 FXR intervention leads to transcriptional improvement of HLC

3.2.1 Transcriptional improvement of FXRi is reproducible across cell lines

To shift the hybrid state towards a hepatic cell fate, an intervention aiming at overexpression of the nuclear receptor FXR with subsequent activation by the FXR agonists chenodeoxycholic acid (CDCA) and GW4064 was conducted and showed improvements in the cell state. Gene expression of several hepatic and hybrid state markers were compared by qRT-PCR, demonstrating that gene expression of FXR itself was reliably increased across all cell lines and protocols by the intervention. Furthermore, reported hepatic FXR downstream targets, such as the transporter proteins BSEP and MRP3 also showed a major increase in gene expression after FXRi across all three iPSC populations (**Figure 3.5**).

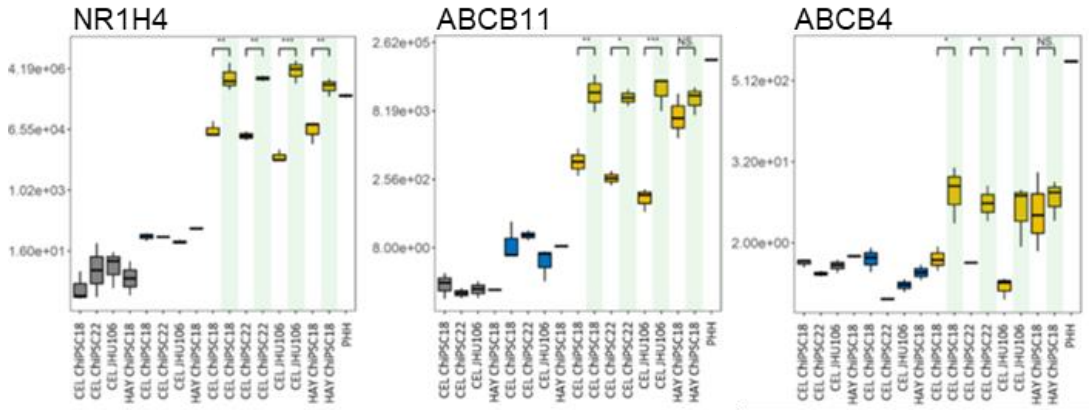


Figure 3.5: Quantitative RT-PCR validation for FXR and FXR target genes. Three genes including FXR and two of FXR-dependent downstream targets (ABCB11, ABCB4) show upregulation after FXRi (highlighted in green) across three different cell lines (ChiPSC18, ChiPSC22, JHU106). N.S.: Not significant, *: p<0.1, **: p<0.01, ***: p<0.001 (Wilcoxon test, unpaired, two-sided).

FXR-independent genes previously defined as markers for successful hepatic differentiation showed a mixed picture. Here, HNF4A remained stable, while FOXA2 showed a slightly positive trend, indicating a possible interaction with FXR. In contrast, ALB was downregulated (not shown). Additionally, expression of genes considered to be critical actors of the intestine-dominated hybrid state of HLC remained stable after the intervention. Here, only the transcription factor ISX showed a mild trend toward downregulation in two out of three tested iPSC derived HLC, while, more importantly, the intestinal master regulator CDX2 remained stable.

Notably, OCT4 and NANOG were consistently downregulated in HLC independent of intervention, in turn indicating a successful exit from pluripotency in all conditions (**Figure 3.6**).

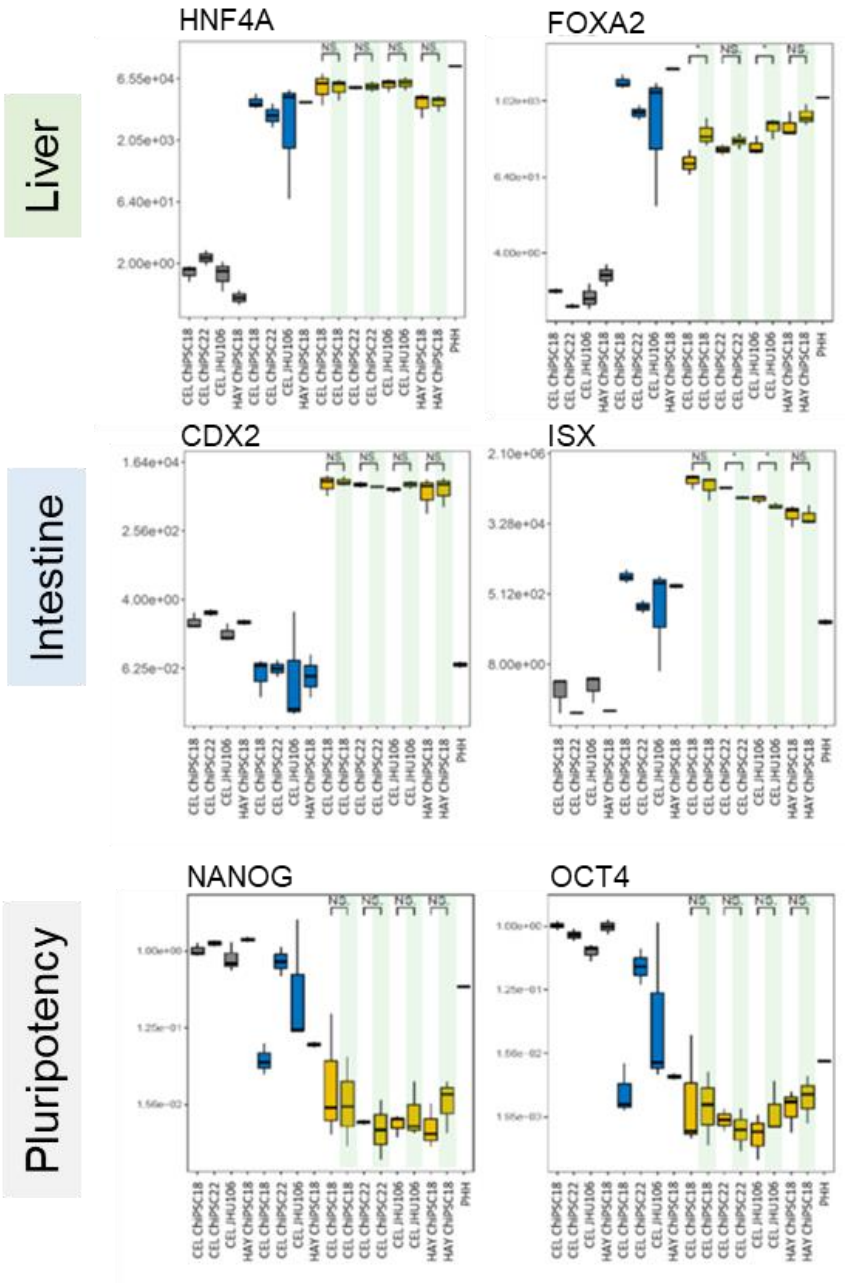


Figure 3.6: Quantitative RT-PCR validation for the example of six marker genes. Relative expression changes of liver marker genes (HNF4A, FOXA2), intestinal marker genes (CDX2, ISX) and pluripotency marker genes (NANOG, OCT4) before and after intervention (FXRi highlighted in green). N.S.: Not significant, *: p<0.1, **: p<0.01, ***: p<0.001 (Wilcoxon test, unpaired, two-sided).

Next, influence of the intervention across different cell lines was evaluated using RNA sequencing. After alignment, differential gene expression analysis was conducted between HLC with and without FXRi, then the resulting DEG were compared between the HLC derived from different iPSC lines. Even though the number of DEG differed, their overlap was more than 5400-fold higher than randomly expected, reflecting the high degree of reproducibility (**Figure 3.7**).

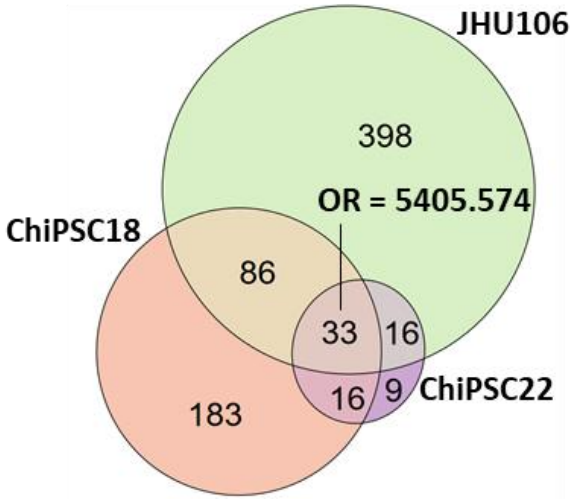


Figure 3.7: Differentially expressed genes due to FXRi across three different cell lines. Shown are the unique and overlapping genes of three iPSC lines (ChiPSC18, ChiPSC22, JHU106) between HLC with and without FXRi with fdr-adjusted p-values lower than 0.001. The overlap was assessed by the overlap ratio (OR), where the randomly expected overlap is 1.0.

Expression plots of selected markers, based on the count data largely corroborated results obtained by qRT-PCR. Here, FXRi caused a consistent increase in gene expression of FXR itself and further known FXR target genes and interaction partners, such as the transmembrane transporter proteins BSEP, MDR2, OATP1B3 and the FXR-dimerization partner SHP (**Figure 3.8**).

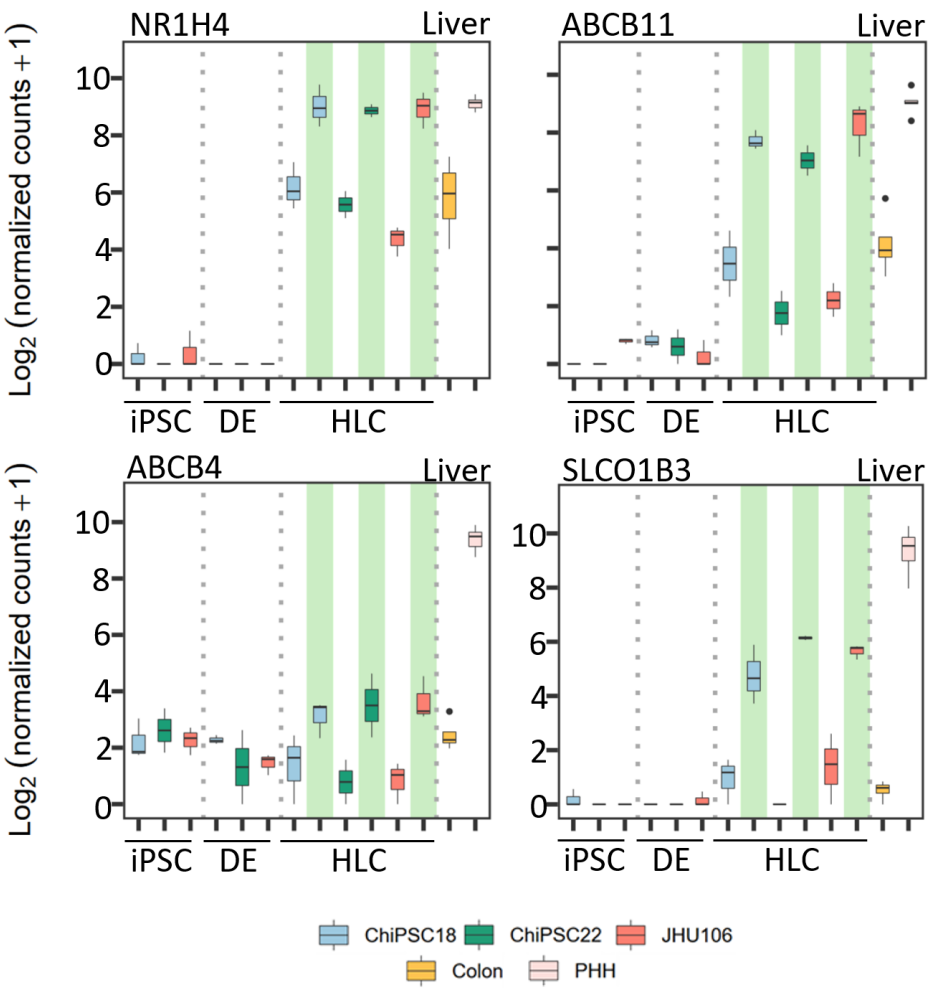


Figure 3.8: Increased gene expression of liver-associated genes after FXRi. Examples include FXR and upregulated FXR-dependent downstream targets (ABCB11, ABCB4, SLCO1B3) due to FXRi (highlighted in green).

The increase of expression levels of liver markers including transcription factors FXR und SHP after FXRi was expectedly mirrored in a decline of the CellNet NIS for those networks. The NIS is a score, which aims to identify transcription factors for the improvement of engineered cells by integrating the expression level of the regulator in the target tissue, the extent of dysregulation of the regulator and its predicted targets in the engineered cell and the number of predicted targets [77]. Even more, several other networks (e.g. ONECUT1) demonstrated a favorable regulation toward a more liver-like status. Surprisingly, this pronounced positive effect on hepatic character of HLC was not reflected in the overall liver GRN score provided by CellNet. By this metric, FXRi showed no significant improvement (**Figure 3.9**).

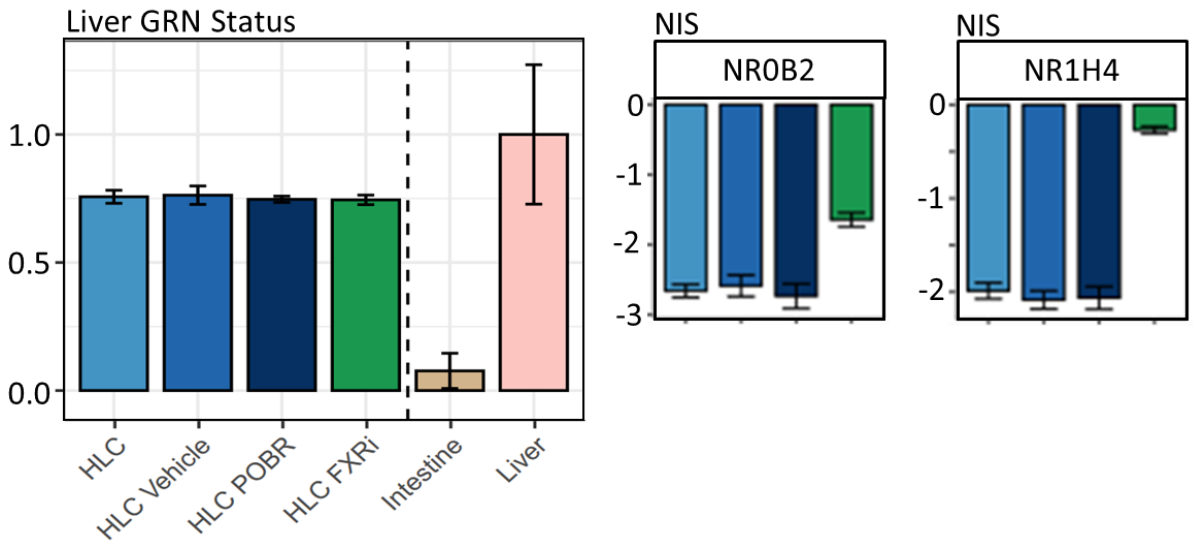


Figure 3.9: CellNet analysis of liver status of HLC. CellNet gene regulatory network (GRN) status (left) and CellNet network influence scores (NIS, right) for liver in HLC derived according to the CEL differentiation protocol with (green) and without FXRi (blue).

The observed suppressing effect of FXRi on intestinal transcription factors CDX2 and ISX was noticeably weaker and corroborated the results obtained by qRT-PCR partially. While there was no significant change in the expression of CDX2, ISX was only slightly downregulated across HLC from three cell lines. On the other hand, the intestine-specific enzyme SI as well as the expression of the brain-associated hybrid gene ALK was significantly downregulated following FXRi (**Figure 3.10**).

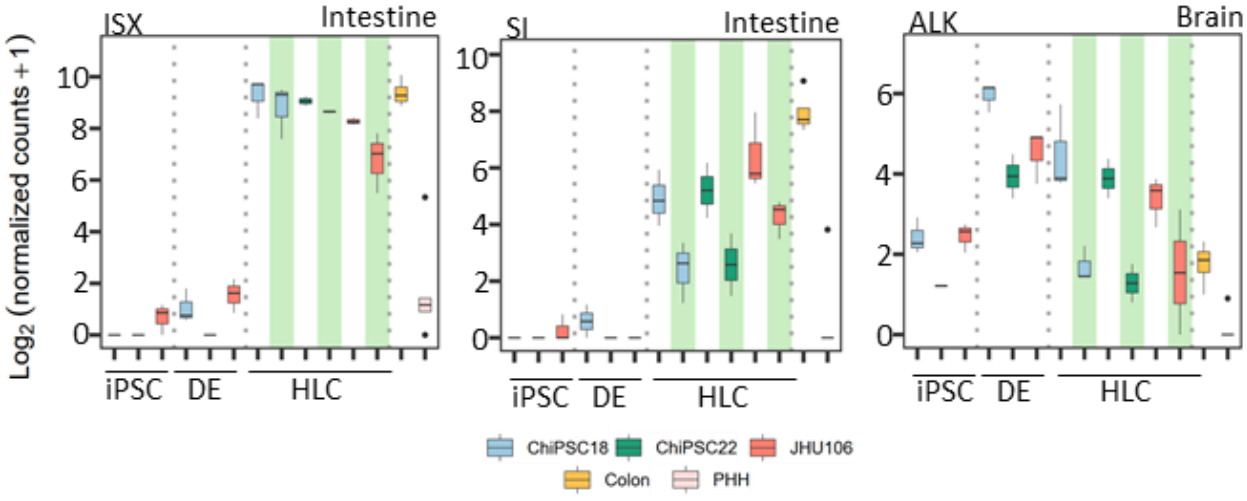


Figure 3.10: Decreased expression of intestine and brain-associated genes after FXRi. Examples of downregulation due to FXRi (highlighted in green) include intestinal genes (ISX, SI) and neuronal receptor tyrosine kinase ALK.

Interestingly, when comparing the expression levels with results obtained by the CellNet algorithm it was found that the network influence score (NIS) of CDX2 and ISX was reduced in HLC after FXRi. Thus, FXRi interfered with these networks in a yet unknown way and the overall reduced intestine GRN status, as provided by CellNet, of these cells could be related to that (**Figure 3.11**).

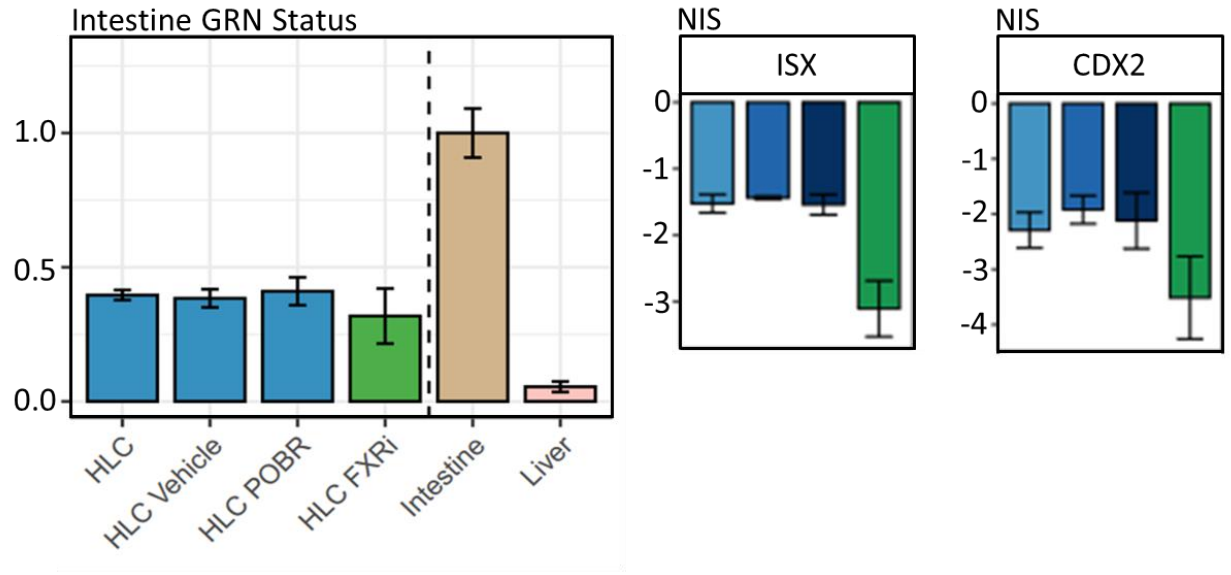


Figure 3.11: CellNet analysis of intestine status of HLC. CellNet gene regulatory network (GRN) status (left) and CellNet network influence scores (NIS, right) for intestine in HLC derived according to the CEL differentiation protocol with (green) and without FXRi (blue).

As shown in further analysis of the FXR liver and colon networks, FXRi had a mostly beneficial influence on the development of liver gene expression, while in contrast, the colon network was impacted negatively. However, in contrast to the CellNet analysis the second effect was less pronounced than the first and differed more across iPSC lines. For an unbiased, genome-wide assessment of the FXRi, the DiPa plot was applied (**Figure 3.12**). Here, the difference of the means of log2FCs of up- and downregulated genes were calculated for each DPG and visualized by a background color scale. For the three tested iPSC lines genes in DPG that should increase to reach similar expression levels compared to PHH were consistently induced (DPG 1, 2, 9, 10) whereas genes in DPG that should decrease to reach similar expression levels compared to PHH were consistently suppressed (DPG 3,4,5,6). The sole exception was DPG 7, where an up- instead of downregulation was observed (**Figure 3.12**).

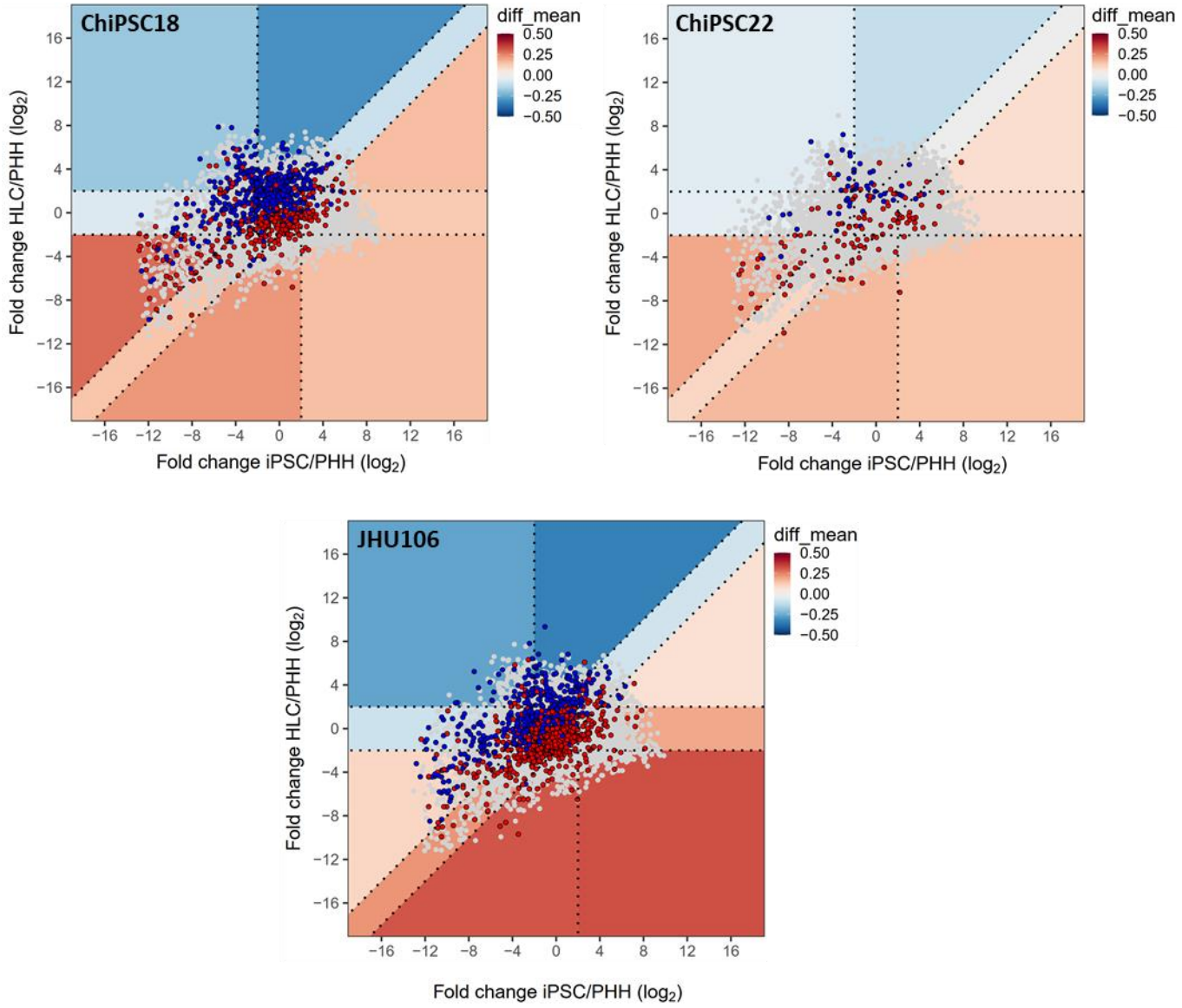


Figure 3.12: Influence of FXRi on HLC differentiation of three different cell lines. Shown are the DiPa plots of HLC from ChiPSC18 (upper left), ChiPSC22 (upper right) and JHU106 (bottom center) cell line differentiated according to the CEL protocol. Red and blue dots indicate genes significantly up- or downregulated by the FXR intervention. The background color indicates mean up- (red) or downregulation (blue) of genes within a DiPa group.

3.2.2 Transcriptional improvement is less pronounced across different protocols

Although protocols to differentiate iPSC and ESC into HLC vary, all of them resort to findings of developmental studies in other vertebrates that have painstakingly put together a still incomplete picture of liver organogenesis for the last decades. Expectedly, many protocols achieve differentiation by means of the same basic principles, yet often use growth factor and small molecule analogues, undefined media components and a varying culture duration which could potentially be drivers of variation between HLC of different origin. Due to feasibility issues, most studies neglect inter-protocol comparisons between HLC of the same donor background. In this context, it was of major interest to reproduce HLC of one iPSC line (ChiPSC18) in a published, non-proprietary protocol and compare them to HLC derived in the CEL protocol before and after FXRi. To that end, differential gene expression analysis and the DiPa supervised clustering approach was applied for HLC derived in the second protocol, termed HAY. Interestingly, it was observed that without FXRi, the gene expression level of FXR in HLC derived via the HAY protocol was already higher than of HLC without FXRi derived in the CEL protocol (**Figure 3.5**). Nonetheless, FXRi solicited a significant response on FXR, ABCB11, ABCB4 and further genes in HLC derived by the HAY protocol, corroborating the results in HLC from three different cell lines derived in the CEL protocol. Given that the hybrid state and the derived FXRi were initially based on HLC differentiated in the CEL protocol, a relatively large overlap of FXR-mediated expression changes with the standard protocol was observed (**Figure 3.13**).

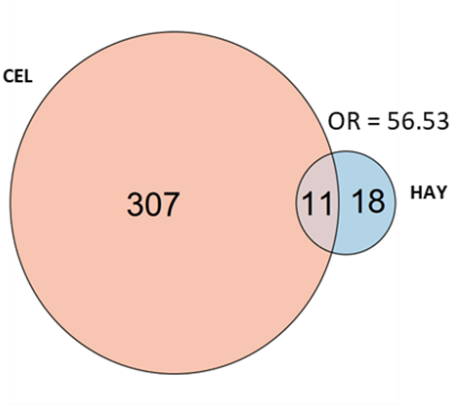


Figure 3.13: Differentially expressed genes due to FXRi across two different protocols. Shown are the unique and overlapping genes due to FXRi of ChiPSC18 HLC differentiated according to the CEL and HAY protocol with fdr-adjusted p-values lower than 0.001. The overlap was assessed by the overlap ratio (OR), where the randomly expected overlap is 1.0.

Nevertheless, DiPa analysis visualizing the global expression changes due to FXRi in HLC derived via the HAY protocol showed several deviations. Conversely, genes of DPG 1 and DPG 2 are mostly downregulated rather than upregulated, indicating a negative influence of FXRi on the differentiation (**Figure 3.14**).

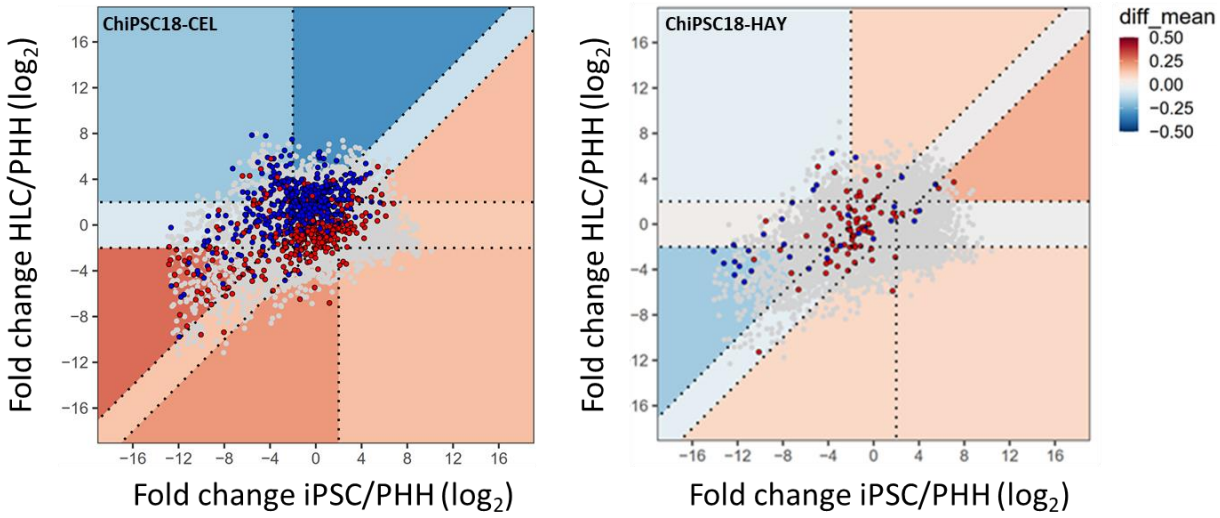


Figure 3.14: Influence of FXRi on HLC differentiation in two different protocols. Shown are the DiPa plots of HLC from ChiPSC18 differentiated according to the CEL protocol (left) and ChiPSC18 differentiated according to the HAY protocol. Red and blue dots indicate genes significantly up- or downregulated by the FXR intervention. The background color indicates mean up- (red) or downregulation (blue) of genes within a DiPa group.

3.3 FXR-related intervention leads to functional maturation of HLC

3.3.1 FXR-related intervention enhances secretory capabilities of HLC

As demonstrated in chapter 3.2, HLC of different origins approach PHH on the transcriptional level. Yet, changes in gene expression are of little consequence when they are not translated to the protein level followed by functional maturation of the HLC. In an initial step to investigate HLC functionality, immunofluorescence stainings for FXR were conducted to show that overexpression resulted in an increase of protein in the nucleus after FXRi (**Figure 3.15**).

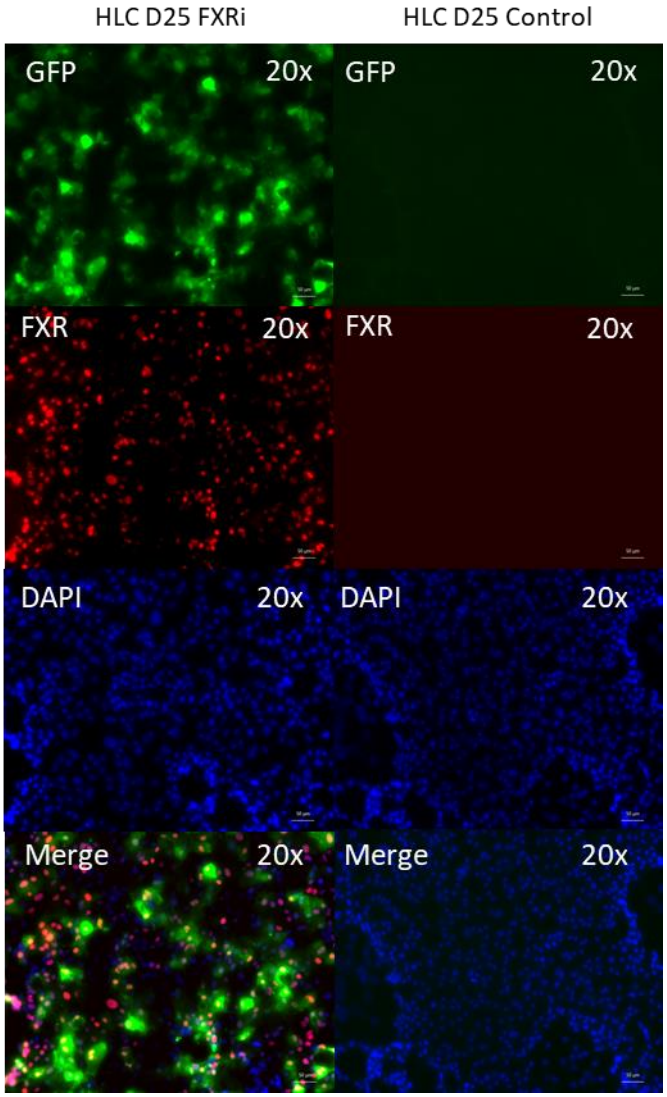


Figure 3.15: Immunostaining shows FXR expression in nuclei of HLC. Immunostaining against FXR on day 25 of differentiation in FXR transduced HLC (HLC D25 FXRi) and control HLC (HLC D25 Control). The expression of EGFP is indicative of cells that successfully integrated the FXR expression construct.

One key function of PHH, which is also of great pharmacological interest is their ability to transport bile acid across membranes. For this, FXR together with other proteins positively regulates expression of certain transmembrane proteins which are located at the canalicular plasma membrane, such as the hepatocyte-specific bile salt export pump (BSEP) protein. As demonstrated, mRNA expression of several of these transport proteins is induced by FXRi. Thus, immunofluorescence stainings were conducted to elucidate if BSEP was induced in HLC after FXRi (**Figure 3.16**).

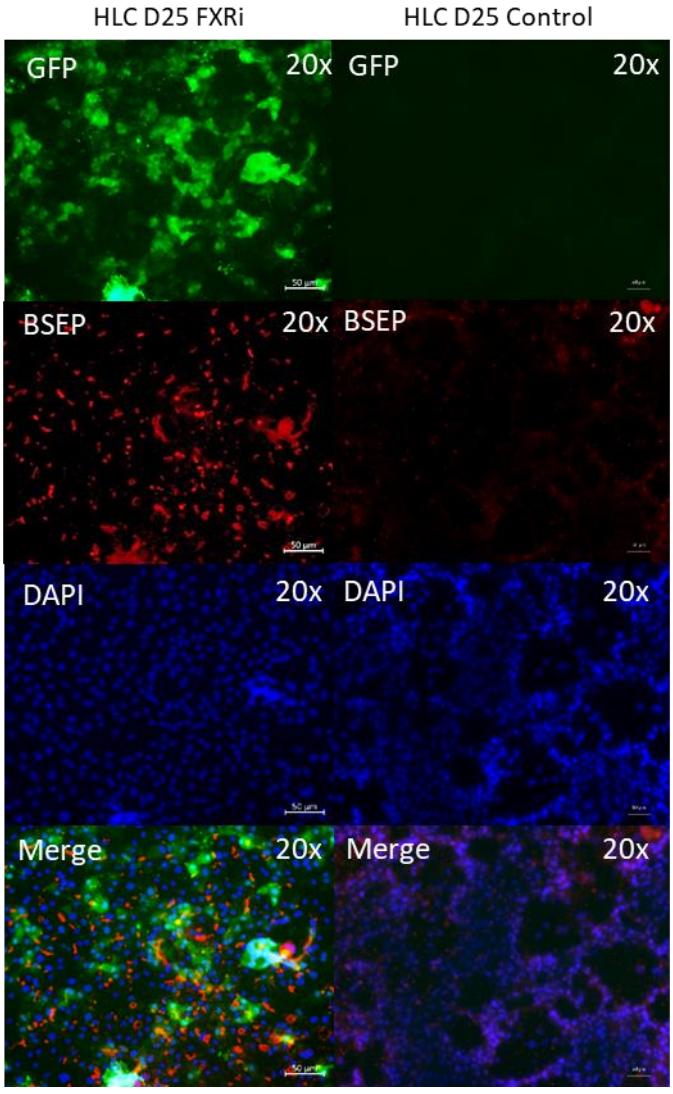


Figure 3.16: Immunostaining shows BSEP expression in nuclei of HLC. Immunostaining against BSEP on day 25 of differentiation in FXR transduced HLC (HLC D25 FXRi) and control HLC (HLC D25 Control). The expression of GFP is indicative of cells that successfully integrated the FXR expression construct.

As compared to the control experiments, HLC with FXRi showed markedly more protein expression of BSEP. Strikingly, protein expression in HLC with FXRi was also polarized at intercellular membrane compartments, indicating correct translocation to the canalicular membrane. Subsequently, it was investigated whether these changes led to functional consequences. For this purpose, CellTracker Green^{TM} dye (5-CMFDA) was used, which upon intracellular transformations turns into an impermeant fluorophore (5-CMF) that needs active transport by BSEP or MDR2 to cross into canalicular structures (FigureCMFDA). After uptake of 5-CMFDA, fluorescence was initially observed in all cells. In control HLC, the cytoplasmic signal was retained longer than in FXRi-treated HLC, which showed remarkably increased secretion rates, in turn leading to a faster decrease of cytoplasmic and an increase of intra-bile canalicular signal (**Figure 3.17**).

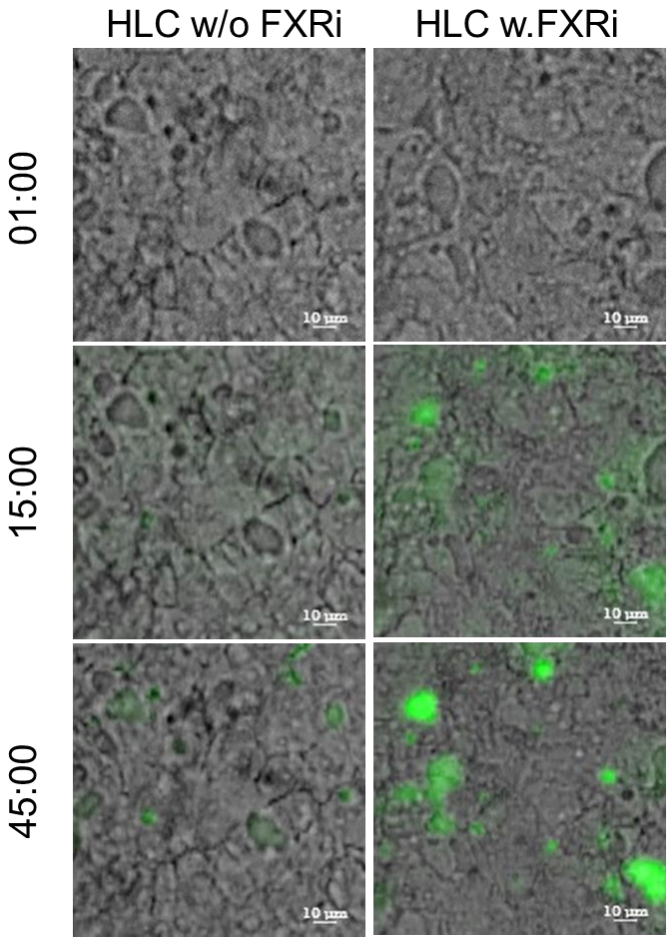


Figure 3.17: FXR activation increases canalicular secretion of HLC. Video stills showing the secretion of green, fluorescent 5-CMF into bile canaliculi after 1, 15 and 45 seconds of exposure of HLC to CMFDA with and without FXR activation (FXR transduction, 1.5 µM GW4064 and 100 µM CDCA).

For quantification, an exponential association function was applied to characterize the increase in canalicular fluorescence by the time constant tau and the amplitude. Here, the time constant tau represents the rate of CMFDA activation and export, whereas the amplitude indicates the amount of activated and exported CMFDA before the transport activity is saturated. The analysis of immune fluorescent experiments and secretion dynamics show that FXRi did not only increase protein expression of bile acid transporter BSEP, but evidently also increase the secretion rate and the amount of secreted CMF per time interval (**Figure 3.18**).

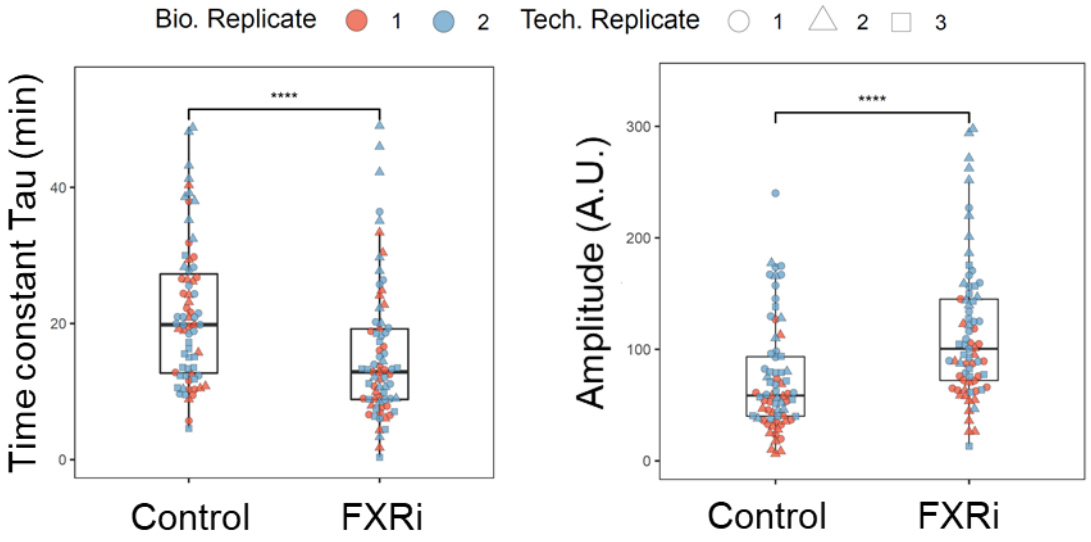


Figure 3.18: Quantification of canalicular fluorescence intensity in HLC. Quantification of the time constant (left) and amplitude (right) of canalicular fluorescence intensity. Each dot represents the data of an individual bile canaliculus; $p<0.0001$ (Wilcoxon test, unpaired, two-sided).

3.3.2 FXR-related intervention enhances lipid droplet formation

Besides secretion and clearing of substances from the cytoplasm, PHH are also involved in storage of lipids and triglycerides into lipid droplets. Here, FXR is known to suppress lipogenesis through downregulation of SREBP. However, in conditions of high amounts of free fatty acids, FXR activity leads to triglyceride accumulation and enlargement of lipid droplets in PHH. This was demonstrated, in an experimental setup, where oleic acid is conjugated with BSA to enable delivery of the fatty acid to cells in culture media. Through subsequent treatment with AdipoRed, the uptake and a potential increase in lipid droplet amount and size can be visualized. Here, untreated PHH showed a consistently low amount of lipid droplets in the cytoplasm. In contrast, PHH treated with an excess of 800 µM of BSA conjugated oleic acid exhibited a major increase in number of lipid droplets as confirmed visually. Expectedly, a control population of PHH treated only with BSA showed no accumulation of lipid droplets (**Figure 3.19**).

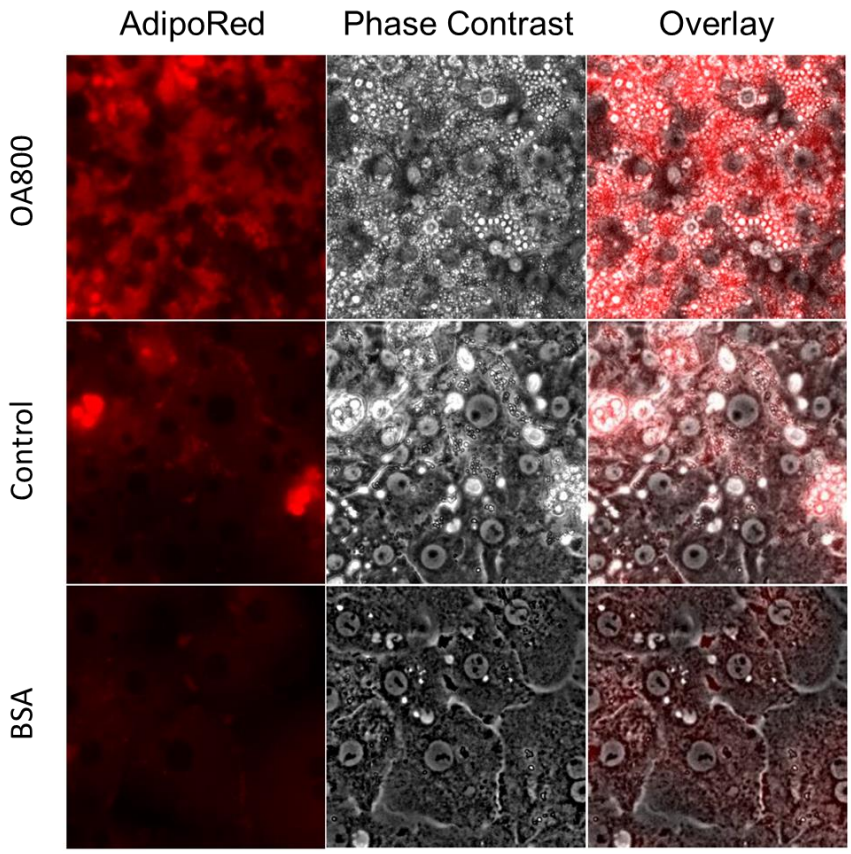


Figure 3.19: PHH form lipid droplets after treatment with oleic acid. Control PHH or PHH treated with 800 µM oleic acid (OA800) or 400 µM BSA were stained with AdipoRed™ after 48 h of incubation.

To examine whether HLC gain the functional capacity to accumulate lipid droplets in an environment of high free fatty acid concentration, the experiment was repeated with HLC with and without FXRi. In HLC without FXRi no major differences in the accumulation of fatty acids could be confirmed visually across the various conditions. In HLC with FXRi however, there was a strong difference in number and size of lipid droplets between HLC untreated and treated with 800 µM oleic acid. Here, the AdipoRed staining exposed a strong resemblance between treated PHH and treated HLC. (**Figure 3.20**).

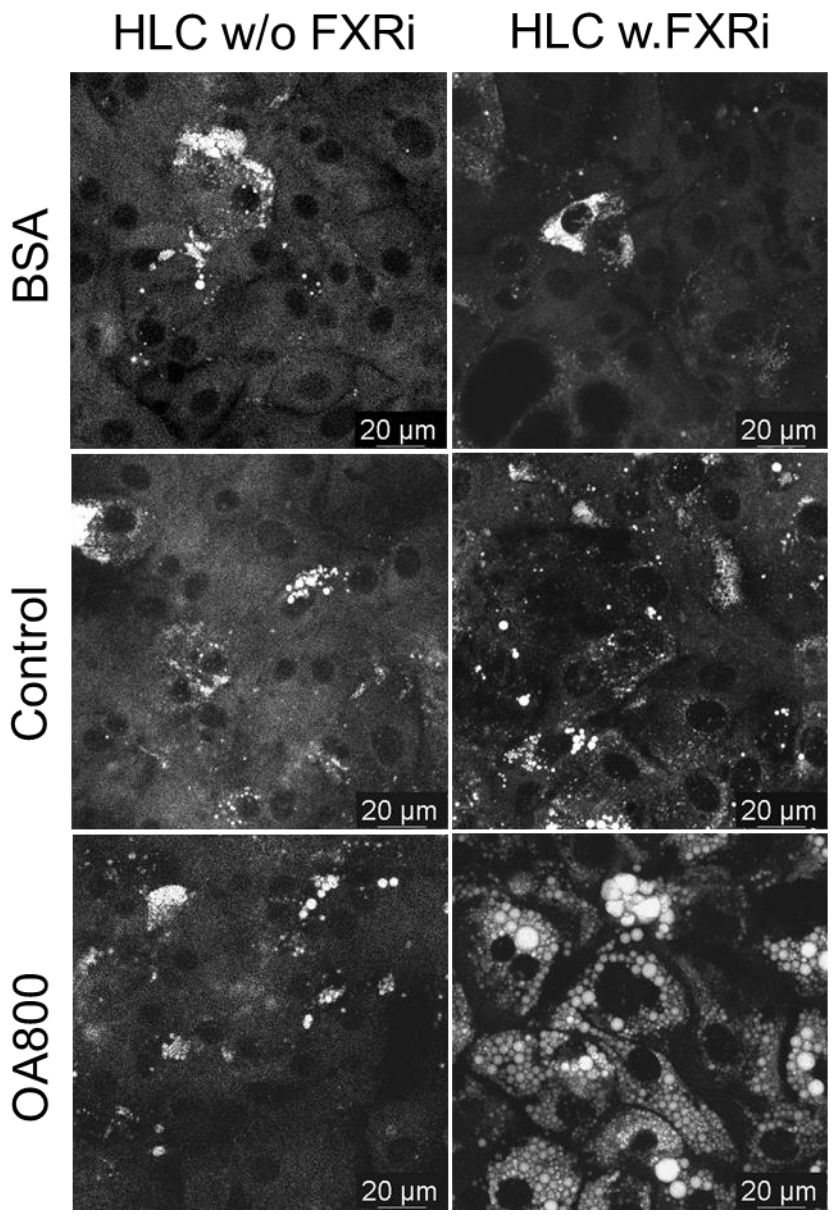


Figure 3.20: HLC increase lipid droplet formation after FXRi. Images AdipoRed staining of HLC with and without FXR, BSA control and oleic acid (OA800) treatment.

The total area of lipid droplets per cell was quantified from AdipoRed stainings of FXRi and control HLC based on the Cellpose machine learning algorithm [70]. By Comparing HLC without FXRi across the various conditions, only a small but statistically significant increase in the area of lipid droplets was found following treatment with 800 µM oleic acid. In contrast, HLC with FXRi showed an enhanced capability to store fatty acids in lipid droplets during high fat treatment compared to control conditions.

Thus, upon FXRi, HLC showed a significant functional improvement, increasing their similarity to PHH (**Figure 3.21**).

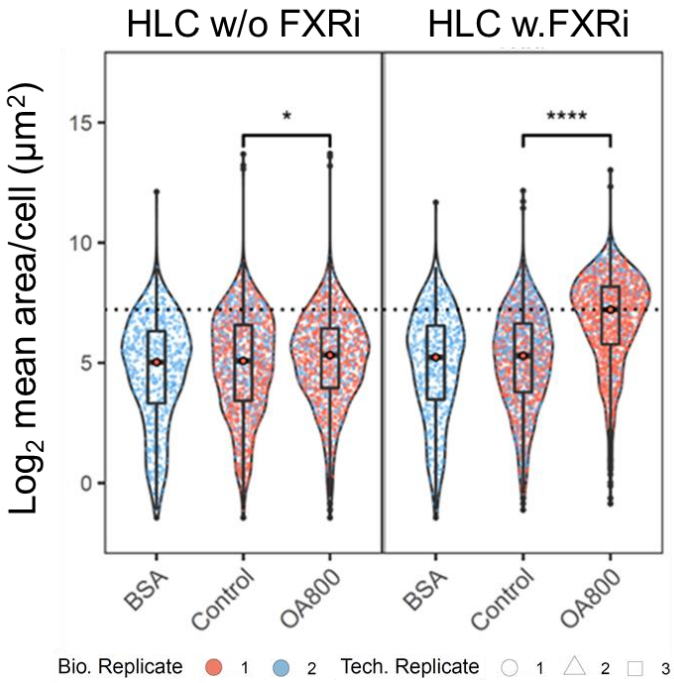


Figure 3.21: Quantification of lipid droplet formation in HLC. Quantification of the total area of lipid droplets per cell in HLC treated with 400 µM bovine serum albumin (BSA), HLC without oleic acid treatment (Control) and HLC treated with 800 µM oleic acid (OA800).

4 Discussion

Primary human hepatocytes remain an important standard tool in clinical research and drug development. Naturally, they are an expensive and limited resource and are subject to dedifferentiation *in vitro*. Thus, there is a need for stem cell-based systems in which HLC cultured according to a differentiation protocol replace PHH. However, since some time it has been established that iPSC-derived HLC suffer from a disturbed gene expression and are widely considered to be immature in comparison to their *in vivo* counterparts. In-depth analysis of transcriptional data demonstrated that HLC are trapped in an artificially induced hybrid state that displays a mixture of several tissue identities but is dominated by hepatic and undesired, mostly intestinal features within one cell. Transcription factors, such as FXR, have been identified to play a pivotal role that can alter the hybrid state on a transcriptional level [1].In this study these findings were substantiated by using a set of different iPSC lines and an additional differentiation protocol to confirm the universal nature of the HLC hybrid state. It was found that the hybrid state is a shared feature in HLC derived from different cell lines and protocols, with some variation depending on the source material and protocol performance. Among these factors, the results exposed that the applied differentiation protocol rather than the genetic background of the iPSC was the main driver of variation in HLC hybrid states. Notably, HAY HLC showed a significant deviation in the number of genes in each DPG compared to CEL HLC. Comparing the number of genes that are upregulated during the differentiation (DPG2 to DPG5), it was found that cells differentiated according to the HAY protocol constantly show lower gene numbers for each DPG, except for insufficiently upregulated genes of DPG2. In contrast, the number of genes with constant low expression (DPG1) is greater in the HAY protocol than in the CEL protocol. This indicates that HAY HLC have a reduced tendency for adverse upregulation as seen in the CEL protocol (DPG4 and DPG5), but on the other side also show less favorable upregulation (DPG3). Thus, the HAY protocol exhibits a generally reduced capacity to induce gene upregulation.

A major concern of cell engineering intervention approaches is the lack of applicability across different iPSC lines. Here, it was demonstrated that FXRi had a major impact on HLC derived from cell lines ChiPSC18 and JHU106 reflected by the number of unique DEG (**Figure 3.7**). The effect of FXRi on the third cell line, ChiPSC22, was less pronounced, reflected by much less unique DEG compared to control experiments. More important however, was the positive effect by FXRi on a core set of genes that was common between all cell lines which included FXR itself and a group of FXR downstream targets. Here, an overlap with an overlap ratio much higher than randomly expected was achieved (**Figure 3.7**). These results were also mirrored in the effect of FXRi on the individual DPG, visualized in differentiation pattern plots for each cell line in which up- and downregulated genes after FXRi were projected based on a change in mean differences before and after intervention (**Figure 3.12**). Here, a clear pattern of downregulation in adverse as well as upregulation in insufficiently expressed DPG was observed, indicating that FXRi has been largely beneficial. Despite the relative small number of genes influenced by FXRi, this positive pattern was also reproduced in HLC derived from ChiPSC22, albeit more modest (**Figure 3.12**). When differentiated according to two different protocols ChiPSC18 reacted differently to FXRi (**Figure 3.13**), showing a protocol-dependent effect of the intervention. Interestingly, HLC differentiated according to the HAY protocol exhibited a stronger expression of FXR already prior to the intervention, which possibly explains the mild effect of FXRi on these cells. Naturally, the objective of this thesis never was to prove that FXRi is a universally applicable intervention path for every HLC but rather that modulation of a single transcription factor suffices to change the balance of the hybrid state in direction to liver. Thus, interventions aiming at nuclear receptors underrepresented in HAY HLC, such as CAR, are promising alternatives for this kind of protocols. The deviation of the HAY protocol from CEL HLC before and after FXRi, therefore only highlights the importance of cross-protocol studies and the need for protocol-tailored intervention strategies for HLC.

This study also aimed to show that FXR intervention not only induces an improvement on the transcriptional level, but increases functional capabilities of HLC as well. Here, it was shown that secretory capacities regulated by FXR and its downstream targets have improved following intervention. Furthermore, lipid storage capacity of HLC increased post intervention, where various HLC exhibited a strong resemblance to PHH treated equally with oleic acid and massively increased lipid storage compared to control HLC.

Although HLC might not yet be a key technology that will soon replace PHH in *in vitro* toxicological and pharmacological studies, it should be noted that, as demonstrated in this thesis, a single transcription factor intervention achieved a functional improvement, robust across several iPSC lines, in crucial drug metabolism-related pathways (Figure 4.1). An open question in the context of functional improvement yet beyond the scope of this thesis is, given the hybrid character of the cells, if HLC also exhibit some intestinal functions and whether these functions decrease in the course of FXR intervention.

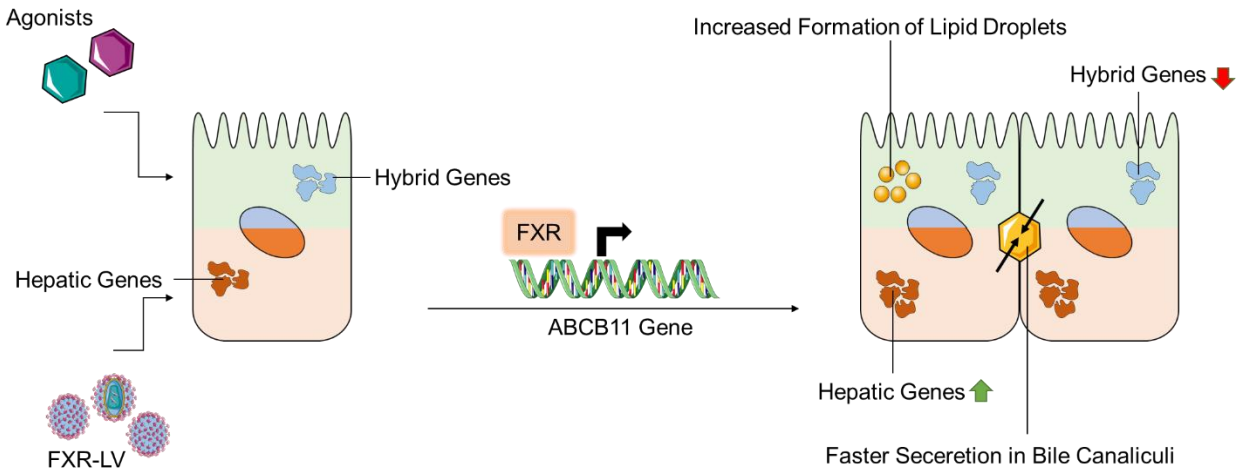


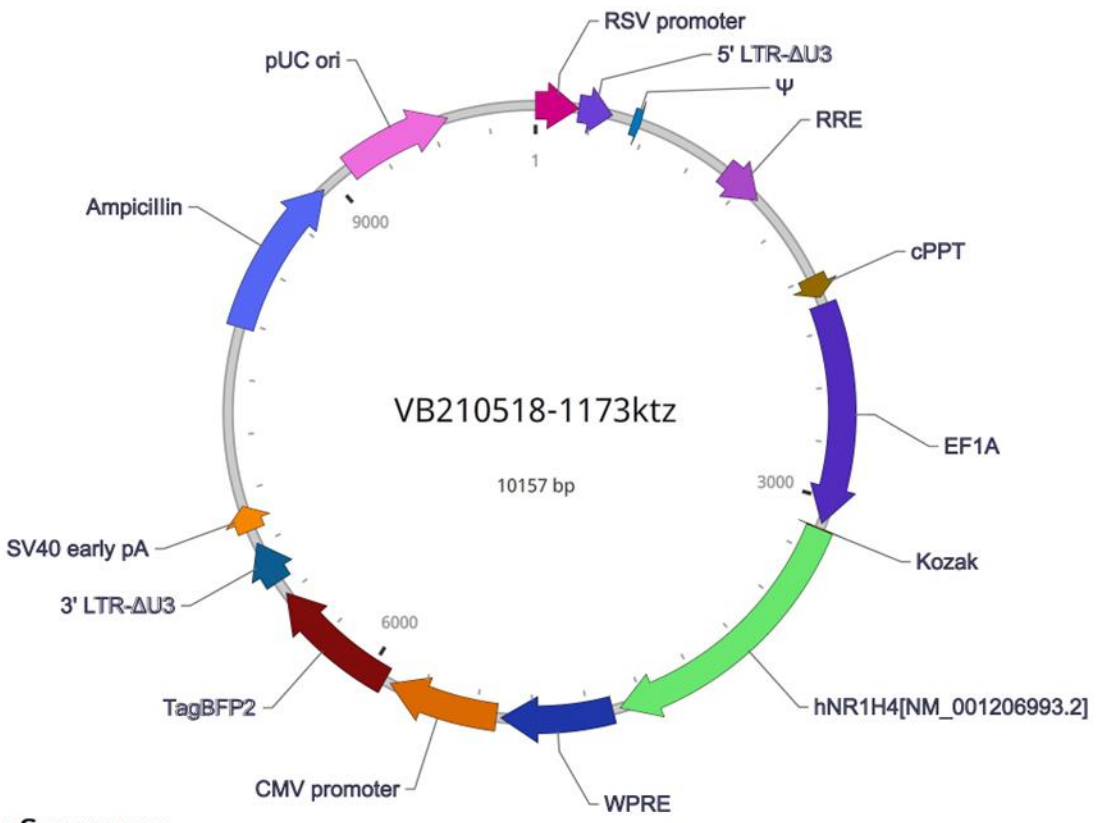
Figure 4.1: Influence of FXR intervention on HLC. Overexpression and activation of nuclear receptor FXR leads to a robust upregulation of hepatic genes and a mild downregulation of undesired hybrid genes. Intervention also improved hepatocyte-associated functionality by increasing the formation of lipid droplets and enabling a faster secretion of compounds into bile canaliculi.

Despite all positive effects achieved by the FXR intervention, expectedly, major differences between HLC and PHH remain. In future, it will therefore be imperative to combine different transcription factor interventions in order to address a wider array of genes and functions of HLC for improvement. However, stable introduction of multiple transcription factors still remains a technological challenge, since viral modes of transduction are limited in number and size of the inserts. Judging by the mild effect of FXR intervention on undesired gene clusters, it is likely that active suppression of hybrid genes such as CDX2 or ISX will be necessary. Here, pilot experiments with shRNA and siRNA have not shown the intended results. More fundamental approaches, in which unfavorable hybrid genes are knocked out using the CRISPR/Cas9 system offer an alternative. Nonetheless, it should be noted that many factors such as CDX2, are not only essential in intestinal tissue maintenance but rather are versatile regulators that perform a wide range of functions at different developmental stages. With all potential intervention strategies and improvements in connection with HLC differentiation approaches, it should not be forgotten that research and development of current iPSC technology is ongoing. New animal studies suggest that OCT4 overexpression is a suboptimal strategy for achieving pluripotency [54]. Although not influencing iPSC gene expression directly, OCT4 seems to interfere with the cells epigenetically, leading to a loss of imprinting and is thus reducing the differentiation potential of the cells significantly [54]. Remarkably, the human iPSC investigated in this study also demonstrated an aberrant epigenetic pattern with high methylation that further increased during the differentiation to HLC [1]. The hypermethylation of gene groups, such as liver specific genes, could be an important factor in the cell potency and therefore needs further investigation. Epigenetic intervention strategies that involve treatment of iPSC or differentiating cells with demethylating agents could be a valuable addition to the aforementioned genetic alterations. Naturally, a study investigating the epigenetic landscape of pluripotent stem cells should also include cells of embryonic origin for comparison. Ultimately, iPSC themselves remain a work in progress and an imperfect technology.

In conclusion, it was demonstrated that FXR overexpression and activation is a robust intervention strategy that allows transcriptional improvement of HLC across various cell lines and protocols. Significantly, these changes also lead to increased protein expression and activity of FXR downstream targets that ultimately brought HLC closer to PHH functionally.

5 Appendix

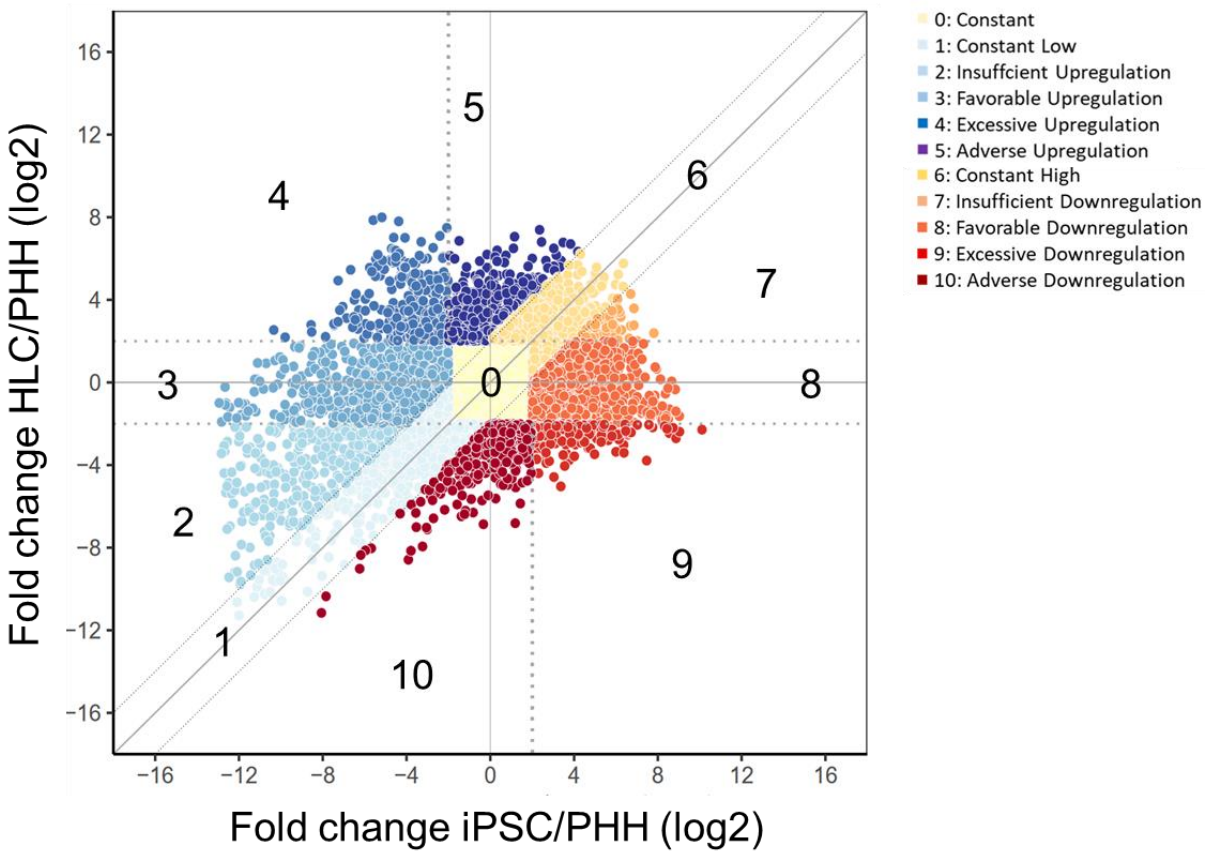
5.1 Supplementary figures



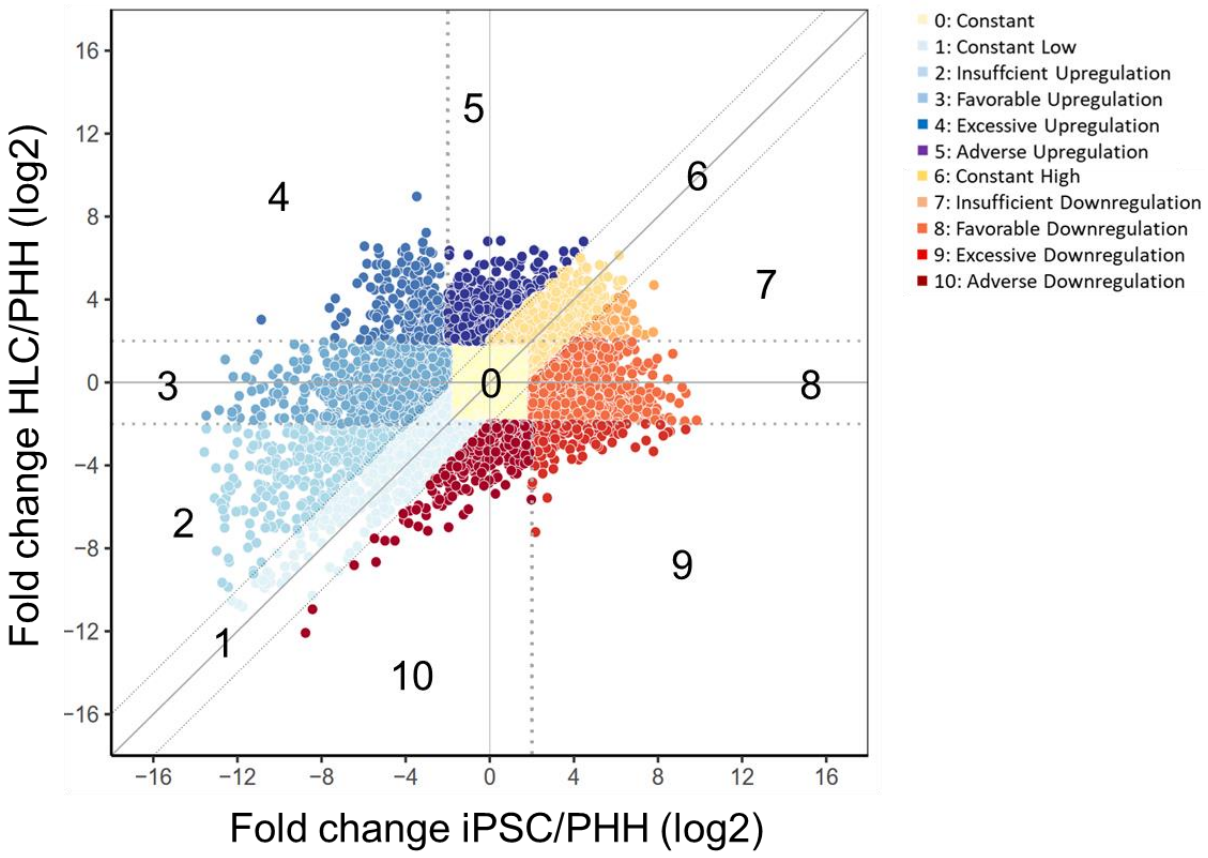
Vector Summary

Vector ID	VB210518-1173ktz
Vector Name	pLV[Exp]-TagBFP2-EF1A>hNR1H4[NM_001206993.2]
Vector Size	10157 bp
Viral Genome Size	6682 bp
Vector Type	Mammalian Gene Expression Lentiviral Vector
Inserted Promoter	EF1A
Inserted ORF	hNR1H4[NM_001206993.2]
Inserted Marker	TagBFP2
Plasmid Copy Number	High
Antibiotic Resistance	Ampicillin
Cloning Host	VB UltraStable (or alternative strain)

Supplementary figure 5.1: Lentiviral vector map. Vector map and additional information of the lentiviral vector used in FXRi experiments.



Supplementary figure 5.2: Differentiation pattern plot of ChiPSC18-derived HLC. Differentiation pattern plot for ChiPSC18-derived HLC differentiated according to the CEL protocol.



Supplementary figure 5.3: Differentiation pattern plot of ChiPSC22-derived HLC. Differentiation pattern plot for ChiPSC22-derived HLC differentiated according to the CEL protocol.

5.2 Supplementary tables

Supplementary tables are available in the digital version of the appendix.

Supplementary table 1: Differentially expressed genes in human induced pluripotent stem cells (iPSC) that were differentiated to definitive endoderm (DE) and subsequently to hepatocyte-like cells (HLC) with and without FXR intervention, compared to primary hepatocytes (PHH) for the cell line and protocol comparison dataset.

Supplementary table 2: Protocol (CEL, HAY) and cell line- (ChiPSC18, ChiPSC22, JHU106) dependent differential gene expression.

Supplementary table 3: Tissue Enrichment Analysis Results for DPG for ChiPSC18-derived HLC.

Supplementary table 4: Tissue Enrichment Analysis Results for DPG for ChiPSC22-derived HLC.

Supplementary table 5: Tissue Enrichment Analysis Results for DPG for JHU106-derived HLC.

Supplementary table 6: Tissue Enrichment Analysis Results for DPG for ChiPSC18-derived HLC in the HAY protocol.

Supplementary table 7: Gene Ontology Results for DPG for ChiPSC18-derived HLC

Supplementary table 8: Gene Ontology Results for DPG for ChiPSC22-derived HLC

Supplementary table 9: Gene Ontology Results for DPG for JHU106-derived HLC

Supplementary table 10: Gene Ontology Results for DPG for ChiPSC18-derived HLC in the HAY protocol.

6 Availability of data and materials

The datasets generated for this study are available in the EGA repository under accession EGAS00001004201.

7 References

[1] Nell P, Kattler K, Feuerborn D, Hellwig B, Rieck A, Salhab A, et al. Identification of an FXR-modulated liver-intestine hybrid state in iPSC-derived hepatocyte-like cells. J Hepatol 2022. https://doi.org/10.1016/j.jhep.2022.07.009.

[2] Blau HM, Brazelton TR, Weimann JM. The evolving concept of a stem cell: Entity or function? Cell 2001;105:829–41. https://doi.org/10.1016/S0092-8674(01)00409-3.

[3] Smith A. A glossary for stem-cell biology. Nature 2006;441:1060–1060. https://doi.org/10.1038/nature04954.

[4] Thomson JA. Embryonic stem cell lines derived from human blastocysts. Science (80-) 1998;282:1145–7. https://doi.org/10.1126/science.282.5391.1145.

[5] Chambers I, Smith A. Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene 2004;23:7150–60. https://doi.org/10.1038/sj.onc.1207930.

[6] Niwa H, Miyazaki JI, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 2000;24:372–6. https://doi.org/10.1038/74199.

[7] Silva J, Smith A. Capturing Pluripotency. Cell 2008;132:532–6. https://doi.org/10.1016/j.cell.2008.02.006.

[8] Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol 2007;9:625–35. https://doi.org/10.1038/ncb1589.

[9] Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell 2011;145:875–89. https://doi.org/10.1016/j.cell.2011.05.017.

[10] Boyer LA, Tong IL, Cole MF, Johnstone SE, Levine SS, Zucker JP, et al. Core

transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005;122:947–56. https://doi.org/10.1016/j.cell.2005.08.020.

[11] Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, et al. Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. Cell 2008;133:1106–17. https://doi.org/10.1016/j.cell.2008.04.043.

[12] Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, et al. Connecting microRNA Genes to the Core Transcriptional Regulatory Circuitry of Embryonic Stem Cells. Cell 2008;134:521–33. https://doi.org/10.1016/j.cell.2008.07.020.

[13] Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 2006;126:663–76. https://doi.org/10.1016/j.cell.2006.07.024.

[14] Jones DL, Wagers AJ. No place like home: Anatomy and function of the stem cell niche. Nat Rev Mol Cell Biol 2008;9:11–21. https://doi.org/10.1038/nrm2319.

[15] Hamey FK, Nestorowa S, Kinston SJ, Kent DG, Wilson NK, Göttgens B. Reconstructing blood stem cell regulatory network models from single-cell molecular profiles. Proc Natl Acad Sci 2017;114:5822–9. https://doi.org/10.1073/pnas.1610609114.

[16] Kraus MRC, Grapin-Botton A. Patterning and shaping the endoderm in vivo and in culture. Curr Opin Genet Dev 2012;22:347–53. https://doi.org/10.1016/j.gde.2012.05.002.

[17] Wells JM, Melton DA. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. Development 2000;127:1563–72. https://doi.org/10.1242/dev.127.8.1563.

[18] Kimura W, Yasugi S, Fukuda K. Regional specification of the endoderm in the early chick embryo. Dev Growth Differ 2007;49:365–72. https://doi.org/10.1111/j.1440-169X.2007.00933.x.

[19] Han L, Chaturvedi P, Kishimoto K, Koike H, Nasr T, Iwasawa K, et al. Single cell transcriptomics identifies a signaling network coordinating endoderm and mesoderm diversification during foregut organogenesis. Nat Commun 2020;11. https://doi.org/10.1038/s41467-020-17968-x.

[20] Tremblay KD, Zaret KS. Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. Dev Biol 2005;280:87–99. https://doi.org/10.1016/j.ydbio.2005.01.003.

[21] Bort R, Signore M, Tremblay K, Barbera JPM, Zaret KS. Hex homeobox gene controls the transition of the endoderm to a pseudostratified, cell emergent epithelium for liver bud development. Dev Biol 2006;290:44–56. https://doi.org/10.1016/j.ydbio.2005.11.006.

[22] Ober EA, Lemaigre FP. Development of the liver: Insights into organ and tissue morphogenesis. J Hepatol 2018;68:1049–62. https://doi.org/10.1016/j.jhep.2018.01.005.

[23] Thompson WL, Takebe T. Human liver model systems in a dish. Dev Growth Differ 2021;63:47–58. https://doi.org/10.1111/dgd.12708.

[24] Kamiya A, Kinoshita T, Ito Y, Matsui T, Morikawa Y, Senba E, et al. Fetal liver development requires a paracrine action of oncostatin M through the gp130 signal transducer. EMBO J 1999;18:2127–36. https://doi.org/10.1093/emboj/18.8.2127.

[25] Miyajima A, Kinoshita T, Tanaka M, Kamiya A, Mukouyama Y, Hara T. Role of oncostatin M in hematopoiesis and liver development. Cytokine Growth Factor Rev 2000;11:177–83. https://doi.org/10.1016/S1359-6101(00)00003-4.

[26] Si-Tayeb K, Lemaigre FP, Duncan SA. Organogenesis and Development of the Liver. Dev Cell 2010;18:175–89. https://doi.org/10.1016/j.devcel.2010.01.011.

[27] Abdel-Misih SRZ, Bloomston M. Liver Anatomy. Surg Clin North Am 2010;90:643–53. https://doi.org/10.1016/j.suc.2010.04.017.

[28] Taub R. Liver regeneration: from myth to mechanism. Nat Rev Mol Cell Biol 2004;5:836–47. https://doi.org/10.1038/nrm1489.

[29] Godoy P, Hewitt NJ, Albrecht U, Andersen ME, Ansari N, Bhattacharya S, et al. Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. vol. 87. 2013. https://doi.org/10.1007/s00204-013-1078-5.

[30] Jungermann K, Katz N. Functional specialization of different hepatocyte populations. Physiol Rev 1989;69:708–64. https://doi.org/10.1152/physrev.1989.69.3.708.

[31] Halpern KB, Shenhav R, Matcovitch-Natan O, Tóth B, Lemze D, Golan M, et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 2017;542:352–6. https://doi.org/10.1038/nature21065.

[32] Michalopoulos G. Liver regeneration. J Cell Physiol 2007;213:286–300. https://doi.org/10.1002/JCP.

[33] Tao Y, Wang M, Chen E, Tang H. Liver Regeneration: Analysis of the Main Relevant Signaling Molecules. Mediators Inflamm 2017;2017. https://doi.org/10.1155/2017/4256352.

[34] Birchmeier W. Orchestrating Wnt signalling for metabolic liver zonation. Nat Cell Biol 2016;18:463–5. https://doi.org/10.1038/ncb3349.

[35] Sun L, Cai J, Gonzalez FJ. The role of farnesoid X receptor in metabolic diseases, and gastrointestinal and liver cancer. Nat Rev Gastroenterol Hepatol 2021;18:335–47. https://doi.org/10.1038/s41575-020-00404-2.

[36] Schulze RJ, Schott MB, Casey CA, Tuma PL, McNiven MA. The cell biology of the hepatocyte: A membrane trafficking machine. J Cell Biol 2019;218:2096–112. https://doi.org/10.1083/jcb.201903090.

[37] Jiao Y, Lu Y, Li XY. Farnesoid X receptor: A master regulator of hepatic triglyceride and glucose homeostasis. Acta Pharmacol Sin 2015;36:44–50. https://doi.org/10.1038/aps.2014.116.

[38] Wu K, Zhao T, Hogstrand C, Xu YC, Ling SC, Chen GH, et al. FXR-mediated inhibition of autophagy contributes to FA-induced TG accumulation and accordingly reduces FA-induced lipotoxicity. Cell Commun Signal 2020;18:1–16. https://doi.org/10.1186/s12964-020-0525-1.

[39] Hingorani AD, Kuan V, Finan C, Kruger FA, Gaulton A, Chopade S, et al. Improving the odds of drug development success through human genomics: modelling study. Sci Rep 2019;9:1–25. https://doi.org/10.1038/s41598-019-54849-w.

[40] Geuens T, van Blitterswijk CA, LaPointe VLS. Overcoming kidney organoid challenges for regenerative medicine. Npj Regen Med 2020;5. https://doi.org/10.1038/s41536-020-0093-4.

[41] Moradi S, Mahdizadeh H, Šarić T, Kim J, Harati J, Shahsavarani H, et al. Research and therapy with induced pluripotent stem cells (iPSCs): Social, legal, and ethical considerations. Stem Cell Res Ther 2019;10:1–13. https://doi.org/10.1186/s13287-019-1455-y.

[42] Godoy P, Schmidt-Heck W, Natarajan K, Lucendo-Villarin B, Szkolnicka D, Asplund A, et al. Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. J Hepatol 2015;63:934–42. https://doi.org/10.1016/j.jhep.2015.05.013.

[43] Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriguen A, et al. Efficient Differentiation of Hepatocytes from Human Embryonic Stem Cells Exhibiting Markers Recapitulating Liver Development In Vivo. Stem Cells 2008;26:894–902. https://doi.org/10.1634/stemcells.2007-0718.

[44] Sullivan GJ, Hay DC, Park I-H, Fletcher J, Hannoun Z, Payne CM, et al. Generation of functional human hepatic endoderm from human induced pluripotent stem cells. Hepatology 2010;51:329–35. https://doi.org/10.1002/hep.23335.

[45] Itzhaki I, Maizels L, Huber I, Zwi-Dantsis L, Caspi O, Winterstern A, et al. Modelling the long QT syndrome with induced pluripotent stem cells. Nature 2011;471:225–30. https://doi.org/10.1038/nature09747.

[46] Tachibana M, Amato P, Sparman M, Gutierrez NM, Tippner-Hedges R, Ma H, et al. Human embryonic stem cells derived by somatic cell nuclear transfer. Cell 2013;153:1228–38. https://doi.org/10.1016/j.cell.2013.05.006.

[47] Camp JG, Sekine K, Gerber T, Loeffler-Wirth H, Binder H, Gac M, et al. Multilineage communication regulates human liver bud development from pluripotency. Nature 2017;546:533–8. https://doi.org/10.1038/nature22796.

[48] Velazquez JJ, LeGraw R, Moghadam F, Tan Y, Kilbourne J, Maggiore JC, et al. Gene Regulatory Network Analysis and Engineering Directs Development and Vascularization of Multilineage Human Liver Organoids. Cell Syst 2020;12:41-55.e11. https://doi.org/10.1016/j.cels.2020.11.002.

[49] Guye P, Ebrahimkhani MR, Kipniss N, Velazquez JJ, Schoenfeld E, Kiani S, et al. Genetically engineering self-organization of human pluripotent stem cells into a liver bud-like tissue using Gata6. Nat Commun 2016;7:1–12. https://doi.org/10.1038/ncomms10243.

[50] Wang X, Yang L, Wang YC, Xu ZR, Feng Y, Zhang J, et al. Comparative analysis of cell lineage differentiation during hepatogenesis in humans and mice at the single-cell transcriptome level. Cell Res 2020. https://doi.org/10.1038/s41422-020-0378-6.

[51] Segal JM, Kent D, Wesche DJ, Ng SS, Serra M, Oulès B, et al. Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors. Nat Commun 2019;10. https://doi.org/10.1038/s41467-019-11266-x.

[52] Hochedlinger K, Yamada Y, Beard C, Jaenisch R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell 2005;121:465–77. https://doi.org/10.1016/j.cell.2005.02.018.

[53] Chen Y, Lüttmann FF, Schoger E, Schöler HR, Zelarayán LC, Kim KP, et al. Reversible reprogramming of cardiomyocytes to a fetal state drives heart regeneration in mice. Science (80-) 2021;373:1537–40. https://doi.org/10.1126/science.abg5159.

[54] Velychko S, Adachi K, Kim KP, Hou Y, MacCarthy CM, Wu G, et al. Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. Cell Stem Cell 2019;25:737-753.e4. https://doi.org/10.1016/j.stem.2019.10.002.

[55] Ouchi R, Togo S, Kimura M, Shinozawa T, Koido M, Koike H, et al. Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids. Cell Metab 2019;30:374-384.e6. https://doi.org/10.1016/j.cmet.2019.05.007.

[56] Pettinato G, Lehoux S, Ramanathan R, Salem MM, He LX, Muse O, et al. Generation of fully functional hepatocyte-like organoids from human induced pluripotent stem cells mixed with Endothelial Cells. Sci Rep 2019;9:1–21. https://doi.org/10.1038/s41598-019-45514-3.

[57] Siller R, Greenhough S, Naumovska E, Sullivan GJ. Small-molecule-driven hepatocyte differentiation of human pluripotent stem cells. Stem Cell Reports 2015;4:939–52. https://doi.org/10.1016/j.stemcr.2015.04.001.

[58] Spence JR, Mayhew CN, Rankin SA, Kuhar MF, Vallance JE, Tolle K, et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature 2011;470:105–10. https://doi.org/10.1038/nature09691.

[59] Jin W, Jiang W. Stepwise differentiation of functional pancreatic β cells from human pluripotent stem cells. Cell Regen 2022;11:1–19. https://doi.org/10.1186/s13619-022-00125-8.

[60] Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, et al. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology 2010;51:297–305. https://doi.org/10.1002/hep.23354.

[61] Czysz K, Minger S, Thomas N. Dmso efficiently down regulates pluripotency genes in human embryonic stem cells during definitive endoderm derivation and increases the proficiency of hepatic differentiation. PLoS One 2015;10:1–16. https://doi.org/10.1371/journal.pone.0117689.

[62] Kamiya A, Kinoshita T, Ito Y, Matsui T, Morikawa Y, Senba E, et al. Fetal liver development requires a paracrine action of oncostatin M through the gp130 signal transducer. EMBO J 1999;18:2127–36. https://doi.org/10.1093/emboj/18.8.2127.

[63] Kamiya A, Kinoshita T, Miyajima A. Oncostatin M and hepatocyte growth factor induce hepatic maturation via distinct signaling pathways. FEBS Lett 2001;492:90–4. https://doi.org/10.1016/S0014-5793(01)02140-8.

[64] Morris SA, Cahan P, Li H, Zhao AM, San Roman AK, Shivdasani RA, et al. Dissecting engineered cell types and enhancing cell fate conversion via Cellnet. Cell 2014;158:889–902. https://doi.org/10.1016/j.cell.2014.07.021.

[65] Cahan P, Li H, Morris SA, Lummertz Da Rocha E, Daley GQ, Collins JJ. CellNet: Network biology applied to stem cell engineering. Cell 2014;158:903–15. https://doi.org/10.1016/j.cell.2014.07.020.

[66] Kong W, Fu YC, Holloway EM, Garipler G, Yang X, Mazzoni EO, et al. Capybara: A computational tool to measure cell identity and fate transitions. Cell Stem Cell 2022;29:635-649.e11. https://doi.org/10.1016/j.stem.2022.03.001.

[67] Wang Y, Alhaque S, Cameron K, Meseguer-Ripolles J, Lucendo-Villarin B, Rashidi H, et al. Defined and Scalable Generation of Hepatocyte-like Cells from Human Pluripotent Stem Cells. J Vis Exp 2017:1–8. https://doi.org/10.3791/55355.

[68] Tiscornia G, Singer O, Verma IM. Production and purification of lentiviral vectors. Nat Protoc 2006;1:241–5. https://doi.org/10.1038/nprot.2006.37.

[69] Gu X, Albrecht W, Edlund K, Kappenberg F, Rahnenführer J, Leist M, et al. Relevance of the incubation period in cytotoxicity testing with primary human hepatocytes. Arch Toxicol 2018;92:3505–15. https://doi.org/10.1007/s00204-018-2302-0.

[70] Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 2021;18:100–6. https://doi.org/10.1038/s41592-020-01018-x.

[71] Van Der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, et

al. Scikit-image: Image processing in python. PeerJ 2014;2014:1–18. https://doi.org/10.7717/peerj.453.

[72] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21. https://doi.org/10.1093/bioinformatics/bts635.

[73] Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-seq data with or without a reference genome. Bioinforma Impact Accurate Quantif Proteomic Genet Anal Res 2014:41–74. https://doi.org/10.1201/b16589.

[74] Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010;11:R106. https://doi.org/10.1186/gb-2010-11-10-r106.

[75] Jain A, Tuteja G. TissueEnrich: Tissue-specific gene enrichment analysis. Bioinformatics 2018:1–2. https://doi.org/10.1093/bioinformatics/bty890.

[76] Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi A J Integr Biol 2012;16:284–7. https://doi.org/10.1089/omi.2011.0118.

[77] Schwab RM, Radley AH, Lo EKW, Cahan P, Tan Y, Kim J. Assessment of engineered cells using CellNet and RNA-seq. Nat Protoc 2017;12:1089–102. https://doi.org/10.1038/nprot.2017.022.

[78] Jain A, Tuteja G. TissueEnrich: Tissue-specific gene enrichment analysis. Bioinformatics 2019;35:1966–7. https://doi.org/10.1093/bioinformatics/bty890.

8 Acknowledgements

First and foremost, I would like to thank **Patrick Nell** for his tireless effort to contribute to this work and his scientific as well as personal advice. We had countless discussions, be it about politics, alternative lucrative business ventures (GenomoSCOPE - patent pending) or just about every day scientific problems – I would not like to miss one of them.

Furthermore, I would like to thank **Prof. Dr. Jan G. Hengstler** for giving me the opportunity to prepare this thesis in his department and for having the first review. I also would like to thank **Prof. Dr. Jörg Rahnenführer** for numerous insightful lessons in statistics and for taking over the second review of this thesis.

Moreover, I would like to thank the team of the **StemNet** group: **Patrick Nell**, **Antonia Thomitzek**, **Andreas Scholtz-Illigens**, **Kathy Belgasmi** and **Nando** for the support and the terrific office atmosphere (whatever office it may be). StemNet only!

For an enormously successful collaboration and hundreds of productive virtual and real-life meetings I would like to extend thanks to **Karolina Edlund** (IfADo), **Prof. Dr. Jörg Rahnenführer**, **Birte Hellwig** (both TU Dortmund), **Prof. Dr. Jörn Walter, Kathrin Kattler**, (both University of Saarland) **Barbara Küppers-Munther** (TakaraBio Europe).

I would also like to extend my gratitude to my friends of the former LivTox Group: **Kathy Belgasmi**, **Dany González Leiva, Sarah Metzler** and **Maiju Myllys**. There were a lot of laughs, a lot of cake and never any troubles!

For countless entertaining lunch breaks and incredible football skills I owe additional gratitude to **Philipp Gabrys**! For her amazing lab managing and her over-the-top dice throwing skills I thank **Karolina Zajac**!

My **family** and **friends** who have always stood by my side. I am grateful and you probably do not even know how much you contributed to this thesis. Thank you for your continuous support in good and in bad times! This work could not have been done without you!

Not least, I would like to thank the people at **IfADo** and especially in the **SysTox** department for creating a wonderful working environment. Thank you, **Rosemarie Marchan, Cristina Cadenas, Karolina Edlund, Nash Vartak, Christoph van Thriel, Ahmed Ghallab Monika Turajski, Wolfram Föllmann, Brigitte Begher-Tibbe, Georgia Günter, Simon Schäfers, Wiebke Albrecht, Tim Brecklinghaus, Gregor Leonhardt, Annika Glotzbach** and many more for your help and support. It was a pleasure getting to know you!

Absolute honorary special thanks goes to my buddy and emotional rescue and work stress relief companion dog **Hans**! You are a true treasure! Never change!

Last, I would like to thank **Maiju Myllys** for her enduring support, her inexhaustible patience and for giving me the strength and the motivation to carry on. I owe you a lot!