Technical University of Munich
TUM School of Medicine and Health

Impact of cellular heterogeneity on tissue function and drug metabolism.

Eva Sofía Sánchez Quant

Complete reprint of the dissertation approved by the TUM School of Medicine and Health of the Technical University of Munich for the award of the

Doktorin der Naturwissenschaften (Dr. rer. nat.)

Chair: Prof. Dr. Heiko Lickert

Examiners:

1. TUM Junior Fellow Dr. Celia Martinez Jimenez
2. Prof. Dr. Maximilian Reichert

The dissertation was submitted to the Technical University of Munich on 6 December 2023 and accepted by the TUM School of Medicine and Health on 13 March 2024.

Table of Contents

Abstract

Single-cell genomics approaches, including low throughput plate-based and high throughput droplet-based approaches, have revealed a wealth of knowledge on cellular heterogeneity in a plethora of tissues and organs, among them the brain and the liver. Low throughput methodologies such as the recently developed snRNA-seq2 methodology allow the capture of rare, sorted cell populations. Here, snRNA-seq2 was successfully applied to explore the transcriptome of sparse oxytocin-expressing hypothalamic neurons of the mouse brain, revealing exclusive gene expression patterns that are mutually exclusive in neuronal subtypes. However, the methodologies to explore the heterogeneity in the regulatory chromatin epigenetic landscape of these sparse populations and their specific characteristics are limited. This thesis presents the initial steps for the development of a plate-based methodology to explore chromatin accessibility in polyploid hepatocytes.

In the liver, further cellular heterogeneity has been described, for instance, regarding the metabolism of endo- and xenobiotic substances within the tissue. This heterogeneity is affected by intrinsic and extrinsic factors, including hepatic steatosis, occurring in chronic non-alcoholic fatty liver disease (NAFLD) or in healthy aging, or the presence of single-nucleotide polymorphisms. How these factors impact the individual hepatocyte functional specialization and their responses toward exposure to a cocktail of drugs is here explored. The assessment of drug efficacy, safety, and toxicity in preclinical early phases of drug development is classically performed in bulk analyses on primary human hepatocytes (PHHs), the gold standard liver *in vitro* model. This thesis aims for the dissection of the metabolic capacity of individual PHHs *in vitro*, and to assess the impact of aging and chronic intracellular fat accumulation on cellular heterogeneity and drug-related metabolism. Here, the phenotyping cocktail approach was applied, and the individual transcriptomic responses were used as a readout of the metabolic capability using a high throughput droplet-based scRNA-seq approach. Four different subgroups of hepatocytes were identified across the four biological replicates and across treatment conditions. These PHH subgroups showed divergent transcriptional responses and metabolic profiles upon incubation with a drug cocktail, intracellular fat accumulation, and both concomitantly.

At the epigenetic level, studies on the upstream molecular machinery governing the transcriptional responses of mammalian hepatocytes at single-cell resolution under fat accumulation and drug exposure are scarce. Herein, this thesis contains an optimization of the experimental conditions for the assessment of the genome-wide chromatin accessibility patterns of PHHs *in vitro* at single-cell resolution using scATAC-seq. Hereby, different chromatin accessibility configurations were found upon treatment conditions such as drug cocktail exposure, intracellular fatty acid accumulation, or both

simultaneously. These different conditions were found to impact the regulatory landscape of individual human hepatocytes, reflecting the observed distinct patterns observed at the transcriptomic level.

Zusammenfassung

Single-cell genomics, einschließlich plattenbasierter Ansätze mit geringem Durchsatz und tropfenbasierter Ansätze mit hohem Durchsatz, haben eine Fülle von Erkenntnissen über die zelluläre Heterogenität in einer Vielzahl von Geweben und Organen, darunter das Gehirn und die Leber, zutage gefördert. Methoden mit geringem Durchsatz, wie die kürzlich entwickelte snRNA-seq2-Methode, ermöglichen die Erfassung seltener, sortierter Zellpopulationen. Hier wurde die snRNA-seq2-Methode erfolgreich eingesetzt, um das Transkriptom von spärlichen Oxytocin-exprimierenden hypothalamischen Neuronen des Mäusegehirns zu erforschen. Es wurden dabei exklusive Genexpressionsmuster aufgedeckt, die sich in neuronalen Subtypen gegenseitig ausschließen. Die Methoden zur Erforschung der Heterogenität in der epigenetischen Chromatinlandschaft dieser spärlichen Populationen und ihrer spezifischen Merkmale sind jedoch begrenzt. In dieser Arbeit werden die ersten Schritte zur Entwicklung einer plattenbasierten Methodik zur Erforschung der Chromatinzugänglichkeit in polyploiden Hepatozyten vorgestellt.

In der Leber wurde eine weitere zelluläre Heterogenität, beispielsweise in Bezug auf den Stoffwechsel von endo- und xenobiotischen Substanzen innerhalb des Gewebes beschrieben. Diese Heterogenität wird durch intrinsische und extrinsische Faktoren beeinflusst. Darunter, die Präsenz von Einzelnukleotid-Polymorphismen oder, die hepatische Steatose, die bei chronischer nichtalkoholischer Fettlebererkrankung (NAFLD) oder beim gesunden Alterungsprozess auftritt. Es wird untersucht, wie sich diese Faktoren auf die funktionelle Spezialisierung der einzelnen Hepatozyten und ihre Reaktion auf die Einwirkung eines Medikamentencocktails auswirken. Die Bewertung der Wirksamkeit, Sicherheit und Toxizität von Arzneimitteln in den frühen präklinischen Phasen der Arzneimittelentwicklung erfolgt klassischerweise in Massenanalysen an primären menschlichen Hepatozyten (PHHs), dem Goldstandard unter den In-vitro-Modellen. Ziel dieser Arbeit ist es, die Stoffwechselkapazität einzelner PHHs in vitro zu untersuchen und die Auswirkungen von Alterung und chronischer intrazellulärer Fettansammlung auf die zelluläre Heterogenität und den arzneimittelbezogenen Stoffwechsel zu bewerten. Hier wurde der Ansatz des Phänotypisierungscocktails angewandt. Die individuellen transkriptomischen Reaktionen wurden als Indikator für die Stoffwechselfähigkeit mit Hilfe eines auf Tröpfchen basierenden scRNA-seq-Ansatzes mit hohem Durchsatz verwendet. Vier verschiedene Untergruppen von Hepatozyten wurden in den vier biologischen Replikaten und unter den verschiedenen

Behandlungsbedingungen identifiziert. Diese PHH-Untergruppen zeigten unterschiedliche Transkriptionsreaktionen und Stoffwechselprofile bei der Inkubation mit einem Medikamentencocktail, bei intrazellulärer Fettakkumulation oder bei gleichzeitiger Verabreichung beider Substanzen.

Auf epigenetischer Ebene gibt es nur wenige Studien über die vorgelagerte molekulare Maschinerie, die die Transkriptionsreaktionen von Säugetierhepatozyten bei Einzelzellauflösung unter Fettakkumulation und Arzneimittelexposition steuert. Diese Arbeit beinhaltet eine Optimierung der experimentellen Bedingungen für die Bewertung der genomweiten Chromatin-Accessibility-Muster von PHHs in vitro bei Einzelzellauflösung mittels scATAC-seq. Dabei wurden unterschiedliche Konfigurationen der Chromatinzugänglichkeit in Abhängigkeit von den Behandlungsbedingungen wie der Exposition gegenüber einem Medikamentencocktail, der intrazellulären Fettsäureanreicherung oder beidem gleichzeitig festgestellt. Es wurde festgestellt, dass sich diese unterschiedlichen Bedingungen auf die regulatorische Landschaft einzelner menschlicher Hepatozyten auswirken, was die beobachteten unterschiedlichen Muster auf transkriptomischer Ebene widerspiegeln.

1. Introduction

In the present thesis, two different single-cell genomics approaches are used to explore the heterogeneity within complex tissues like the brain and the liver. The heterogeneity among hepatocytes in the liver is dissected by single-cell genomics approaches at the transcriptomic level using single-cell RNA sequencing (scRNA-seq), and at the epigenomic level applying single-cell Assay for Transposase-Accessible Chromatin (scATAC-seq). Thereby, the transcriptomic and epigenomic profile of a seemingly homogeneous population of primary human hepatocytes (PHHs) *in vitro* is interrogated. Moreover, the capacity to metabolize a cocktail of drugs by individual hepatocytes is investigated to deeply characterize their transcriptomic responses. In addition, the impact of intracellular fat accumulation in PHHs *in vitro* on the individual hepatocyte metabolism, and on their capacity to metabolize a cocktail of drugs is evaluated at single-cell resolution. The aims for this thesis are described in section 1.8 "Hypotheses and aims of this thesis" at the end of the Introduction section.

1.1 Single-cell genomics technological approaches: comparison of strategies and relevance of its application to dissect cellular heterogeneity

1.1.1 Single-cell RNA sequencing (scRNA-seq)

Two different scRNA-seq approaches are applied in each of the chapters in this thesis depending on the research question, the sample origin and conservation method, and the throughput– such as the exploration of i) the transcriptomic profile of rare cellular subpopulations; or ii) the characterization of responses within a cell population. First, the exploration of a rare population in the murine brain such as oxytocin-positive hypothalamic neurons in response to control or high-fat/high-sugar diet using flash-frozen tissue samples is performed employing a low-throughput plate-based approach. Second, the characterization of the transcriptional profile among a seemingly homogeneous population of PHHs *in vitro* in response to intracellular fat accumulation, exposure to several drugs simultaneously, or both concomitantly is performed using a high-throughput droplet-based approach (10X Genomics).

In the past years, advances in single-cell genomics technologies have unraveled a wealth of knowledge obtained from analyzing the gene expression profile of single cells in healthy vs. diseased statuses of several organs and systems [2-9]. Cell atlases of different organs and tissues have been developed to gain a comprehensive view of mouse [10] and human, culminating in the *Tabula muris* atlas of mice

tissues, the Human cell atlas [11, 12] and the *Tabula sapiens* [13]. Naturally, a plethora of scRNA-sequencing methods and approaches have emerged fashioned to accommodate the quickly growing research questions and experimental designs. This technology has enabled the interrogation of complex tissues, and cell types and the identification of groups of cells with similar transcriptomic profiles and dynamic states, which are generally associated with similar functional processes [14-17].

Firstly, the isolation of viable single cells can be achieved through different experimental methodologies, depending on the starting material, namely tissue biopsies, primary cells, or cell lines and their preservation: fresh, fixed, or flash-frozen [18-20]. In order to singly dissociate cells, the main commonly used strategies are fluorescent-activated cell sorting (FACS), magnetic-activated cell sorting (MACS), laser capture microdissection (LCM), and microfluidics [21-24]. Alternatively, the isolation of single nuclei can prove beneficial when interested in previously flash-frozen and archived samples, such as those in biobanks [19]; or when the tissue's anatomical complexity hinders the readily isolation of intact cells, for instance from adult brain tissue [25]. The harsh enzymatic dissociation damages RNA integrity, biases recovered cell type proportions and is successful only for young tissues [26, 27].

Two main strategies stemmed for scRNA-seq, diverging in the means whereby transcripts are captured, the number of cells or nuclei analyzed in one run or throughput, the barcoding technique and the sequencing method (**Figure 1**) [28, 29]. Based on their captured transcript coverage, the current single-cell approaches can be divided into full-length transcript capture and tag-based methods capturing polyadenylated mRNA molecules on the 3' or 5'-end transcripts [30, 31]. Isoforms or gene fusions can be studied using full-length transcript methods, many based on the SMART-seq2 [31] technology [30], whereas 3'-end-based methods can provide an aggregate view of the transcriptional heterogeneity of the same cell population, such as 10X Genomics Chromium [28]. Classical droplet-based approaches such as Drop-seq [32], Seq-Well [33], 10X Genomics Chromium [34], inDrop [35], and DroNC-seq [26] are restricted to the capture of one end of the transcript, reducing the ability to unambiguously align reads to a transcript, and detect different gene isoforms, easily performed applying full-length transcript methods such as SMART-seq2 [36].

Another key difference between methodologies relies on the throughput, whereby plate-based methods yield a comparatively low throughput, while droplet-base approaches provide higher output at a generally relatively lower cost per cell, trading off sensitivity [28]. Therefore, plate-based methodologies are more suitable for studies of rare cell types or subpopulations that are scarce in the tissue of origin, for instance, hypothalamic neurons in murine brain tissue [37]. Additionally, plate-based methods are performed on cells sorted into microtiter plates, where the conversion to cDNA occurs per cell with the aid of microfluidic robots [38]. Recently, the snRNA-seq2 method based on

the SMARTer chemistry from Takara was developed in the Martinez-Jimenez lab to explore a starred feature of the liver, polyploidization [19, 39].

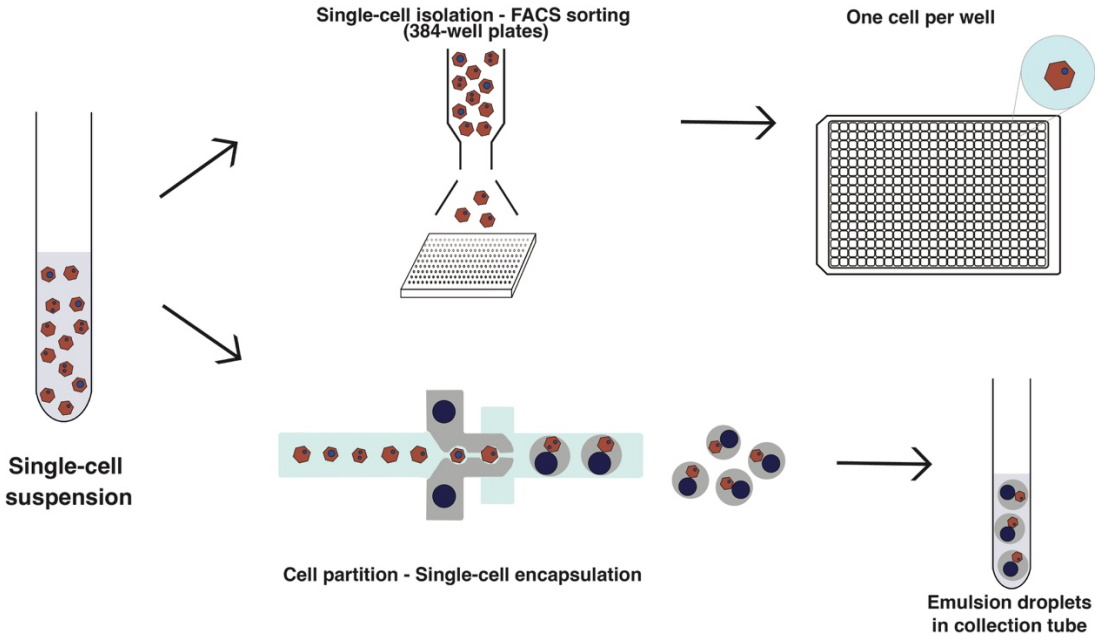


Figure 1. Schematic representation of the workflow for single-cell genomics using plate-based (top) or droplet-based (bottom) approaches. *Plate-based methods rely on the deposition of single cells isolated typically by FACS sorting into each well of a PCR microtiter plate filled with lysis buffer. On the other hand, droplet-based methodologies use oligo-barcoded coated beads encapsulated with cells in emulsion droplets of oil.*

Conversely, droplet-based approaches such as 10X Genomics, encapsulate cells in oil or gel droplets together with the cell barcode and the unique molecular identifier (UMI), followed by droplet lysis and cDNA synthesis [40]. The usage of random nucleotide sequences (UMIs) associated to a unique single molecule can be used for the quantification of gene expression levels and to correct for amplification bias, helping accurate quantification of transcripts [30, 41].

1.1.2 Single-cell ATAC-sequencing (scATAC-seq)

In the present thesis, the preliminary assessment of the impact of these factors on PHHs' chromatin landscape was performed. Scarce single-cell transcriptomic analyses of liver polyploidization have been done [19, 42], yet the epigenetic landscape characterizing polyploid cells in the liver has not been explored [43]. Therefore, in the present thesis, the development of a plate-based methodology for the analysis of the chromatin state of polyploid hepatocytes is one of the objectives.

Gene expression is tightly regulated by genomic and epigenomic molecular events that define the downstream transcriptomic signature of cell types, subtypes, and states. Single-cell studies of the genome-wide chromatin architecture and accessibility landscape have revealed differences in cell types that differ in healthy and diseased statuses [44-50]. The upstream molecular events governing gene expression can be explored in order to gain a broader understanding of the underpinnings regulating processes and functional outcomes, cell identities, and cell types or cell subpopulations [51, 52].

The most widely used strategy to measure the level to which specific regions of chromatin are accessible to regulatory factors is the genome-wide chromatin accessibility landscape assessment through the Accessible Transposase Chromatin Accessibly assay (ATAC-seq), initially developed by Buenrostro *et al.* in 2013 [53]. However, the chromatin state is specifically and dynamically regulated per cell-type [54, 55], hence the methodology was optimized to dissect this heterogeneity at single-cell resolution in 2015 [44, 56]. This methodology relies on the Tn5 hyperactive transposase enzyme [57, 58] entering the permeabilized nuclei to fragment and tag open chromatin regions inserting sequencing adaptors in a process named tagmentation. Open regions are defined by the nucleosome and transcription factor-free areas [53]. ATAC-seq libraries sequenced in paired-end reads provide information about nucleosome positioning and packing, since the insert size distribution of sequenced fragments from human chromatin show a clear periodicity of approximately 150 to 200 bp, with 147 bp empirically tested to be the length of DNA wrapped around one nucleosome [53, 59].

As occurring with scRNA-seq, plate-based [60], nano well-based [61], microfluidics devices (ICELL8 from Takara) [44] and droplet-based methodologies [47, 62], as for instance the 10X Genomics platform [50] have emerged to shelter both research and biological demands. The latter methodology has become the most broadly used method due to its high throughput, the low cost per cell, the capture efficiency and the flexibility it offers [63]. Common to all is the need for the isolation of intact, singly dissociated nuclei. This is the initial and crucial step in the workflow, and the optimization of the lysis conditions and incubation time is key for successful high-quality dataset obtention. To achieve that, mechanical dissociation using douncer homogenizers and lysis buffers are combined in order to obtain a clean single-nuclei suspension, as in the improved Omni-ATAC protocol by Corces *et al.* [64, 65]. The composition of the lysis buffer is conventionally made of nonionic detergents and different concentrations of salts with the purpose of lysing the cellular membrane and permeabilizing the intact nuclear membrane, and needs to be optimized for the tissue or sample of interest [48, 66-68]. Subsequently, the transposition reaction can be performed in bulk, prior to the obtention of single nuclei, as in the pipeline published by Chen *et al*. [60]. Alternatively, the tagmentation can be

performed individually per nuclei after single nuclei are separated into single wells [44]. The selection of the tagmentation strategy and the general methodology relies on the research question.

1.2 Liver and brain connection

The liver is the largest solid organ in the body responsible for vital functions such as the metabolism of endogenous substances (endobiotic) and exogenous substances (xenobiotics) [69, 70], energy homeostasis maintenance, detoxification, coagulation, nutrient transport, or immune response [71, 72]. The neural connection between the liver and the paraventricular nucleus (PVN) of the brain is the backbone machinery to maintain systemic glucose homeostasis [73]. Oxytocin-expressing paraventricular hypothalamic neurons (PVN^{OT} neurons) project directly to the liver and adipose tissue [73, 74], and respond to afferent signals from the gut, one of them being the release of the peptide cholecystokinin (CCK), being major regulators of whole-body energy homeostasis [75]. Nonetheless, the molecular mechanisms by which PVN^{OT} neurons orchestrate gut-to-brain feeding control remain unclear. Recently, scRNA-seq enabled the identification of 62 neuronal and 11 non-neuronal subclasses showing distinct transcriptional signatures in the adult mouse hypothalamus, and up to 24 distinct neuronal types in the median eminence of the arcuate nucleus (ARH-ME) [76-79] as well as the importance of neuronal subtypes of Pro-opiomelanocortin (POMC) anorexigenic neurons in the ARH that signal to the PVN in response to leptin to control feeding and satiety homeostatic signals [80, 81]. Moreover, it is known that chronic exposure to different hypercaloric diets leading to obesity in individuals causes defects in the gut-brain communication. Especially, the suppression of food intake that CCK has is diminished upon high-fat feeding and is associated with attenuated neuronal activation in the PVN [82].

1.3 Liver structure, metabolic function, and drug metabolism

In the liver, the different functions are coordinated and performed by the different hepatic cell types. Among those, constituting the non-parenchymal cell compartment are hepatic stellate cells (HSCs), Kupffer cells (KCs), liver sinusoidal endothelial cells (LSECs), lymphocytes and biliary cells [83]. Up to twenty different cell types have been described in the human liver, among which hepatocytes, endothelial cells, cholangiocytes, HSCs, B-cells, conventional and non-conventional T cells, NK-like cells, and distinct intrahepatic monocyte/macrophage populations were defined by MacParland *et al.*

in 2018 [84]. The main metabolic functions are predominantly performed by hepatocytes, the major cell type of the liver, making up the parenchyma comprising nearly 70% of the liver mass and total cell population in humans, and 50-60% in mice [69, 85-87]. A particularity that hepatocytes have is polyploidy [43, 88-90]. Polyploidy or whole-genome duplication is the inheritable condition to possess more than two complete sets of chromosomes [91]. After birth, all hepatocytes are diploid (2n), and the polyploidization (4n, 8n, 16n) process occurs gradually with postnatal development [92, 93]. In adulthood, up to 90% of hepatocytes are polyploid in rodents [92, 94, 95], and approximately 30 - 40% in humans [89, 96-98].

Spatially, the liver is made up of smaller organizational units named lobules, composed in their structure by a hexagonal shape of around 15 concentric layers of cells [99]. Nutrients, oxygen, hormones, and xenobiotic substances proceed from the intestine and enter the lobule at the peripheral area, flowing towards the center of the hexagon, where the draining central vein (CV) is located. On the contrary, bile flows from the center of the lobule, draining into the peripheral portal bile duct [100-103]. This particular spatial structure prompts a gradient in the concentration of the affluent nutrients and hormones [103], resulting in cellular specialization, known as metabolic zonation [104]. This entails differential gene expression, division of labor, and functional heterogeneity within the liver lobule [2, 5, 6, 17, 84, 105, 106]. Recently, single-cell RNA sequencing, immunohistochemical techniques and spatial reconstruction have revealed the heterogeneity hidden in liver tissue among different cell types and within the same one in health and diseased statuses. A comprehensive characterization of the human and murine liver cell type composition and their communication healthy tissue has been done [2, 5, 6, 8, 17, 84, 105, 107, 108]. In the murine liver, Halpern *et al.* were the first ones to describe liver heterogeneity across the lobule and functionality among cell types [5, 6]. In human tissue, Aizarani *et al.* constructed a liver cell atlas from nine liver donors, describing the heterogeneity hidden in the different cell type populations and proposing an epithelial progenitor in liver development [2].

This has not only been demonstrated at the transcriptomic level but also at the gene regulatory level by characterizing the epigenetic landscape of single cells. For instance, a study by Brosch *et al.* demonstrated that hepatocytes located in different zones carry different epigenetic marks such as DNA methylation patterns [109]. Nevertheless, the chromatin accessibility landscape of individual cells in the liver has not been extensively studied yet. Cusanovich *et al.* made a comprehensive scATAC-seq atlas of 13 tissues, where hepatocytes clustered separately based on their chromatin pattern [45]. Thereafter, Chen *et al.* published in 2020 the first study from fresh murine hepatocytes at single-cell resolution using 10X Genomics applications, where they identified functional diversification of

hepatocytes during liver regeneration based on their differential chromatin accessibility configurations [48]. Also, single-cell bisulfite sequencing has revealed heterogeneous DNA-modification states among single hepatocytes [110]. Most recently, a study using nuclei from mouse livers showed differential chromatin accessibility in young vs. old livers, and while chromatin accessibility patterns were able to separate cells by age, this was not possible from the transcriptomic data, and also the correlation between the epigenomic and the transcriptomic level was generally not optimal [111]. Yet, they report that age is a relevant factor for explaining transcriptional cell-to-cell variation. However, a deep characterization of the metabolism of human hepatocytes at the transcriptomic and epigenomic level has not yet been deeply performed. Also, not in response to exposure to several drug treatments or xenobiotic metabolism, or in response to intracellular fat accumulation.

The metabolism of endobiotic substances in the liver comprises carbohydrates, lipids, bile acids and sterols, protein synthesis, and urea metabolism. Included among xenobiotics are drugs, toxicants, environmental toxicants and pollutants but also flavorings, cosmetics or food additives [112, 113]. Three main phases compose the metabolism of xenobiotics in vivo: Phase I solubilization reactions, phase II conjugation reactions and phase III or transmembrane transport [114]. *In vivo*, the biotransformation of a xenobiotic substance (pharmacokinetics) in an organism goes through four different phases: Absorption, Distribution, Metabolism and Excretion (ADME) [115]. However, *in vitro*, the absorption and metabolism phases are not considerable, and the clearance of administered drugs occurs in three main phases: Phase I, II and III metabolism (**Figure 2**).

1.3.1 Phase I drug metabolism

Phase I metabolism consists of reduction, oxidation or hydrolysis reactions, which convert lipophilic drugs into polar compounds by adding or exposing moieties like -NH_2 or OH to increase their solubilization [114]. Their reactions also produce active metabolites of the drug, resulting in the activation of prodrugs into their active compound, with therapeutic effect [114]. The cytochrome P450 (CYP450) superfamily of monooxygenase enzymes catalyze these biotransformation reactions of drugs and other lipophilic compounds [116-118]. They are expressed as membrane-bound proteins mostly found in the endoplasmic reticulum of hepatocytes [119]. The subfamilies 1, 2, and 3 are the most highly expressed and are responsible for the metabolism of most of the drugs and other xenobiotic chemicals. Precisely, the isoforms CYP1A2, 2C9, 2D6, 2C19 and 3A4 are responsible for the metabolism of approximately 70 to 80% of the drugs currently in the market [117, 120, 121]. The most

abundantly expressed in the healthy human liver are CYPs 3A4, 2C9, 2C8, 2E1 and 1A2, whereas 2A6, 2D6, 2B6, 2C19 and 3A5 are less abundant [117, 122, 123]. This superfamily of enzymes is encoded by 57 genes in humans [124]. These are inducible, meaning their expression is stimulated in the presence of their substrate [117, 119, 125-127].

There is certain overlapping functionality and substrate specificity among them, but many drugs are metabolized at clinically relevant concentrations by only one isoform of the enzyme [128]. For instance, the main route for the metabolism of caffeine transforming *N*-3-demethylation to paraxanthine (1,7-dimethylxanthine or 17X), accounting for approximately 80%, is carried out by the *CYP1A2* isoform [129]. Though, a minor percentage of it is metabolized by other enzyme isoforms such as CYP2C8/9 and CYP3A4. Similarly, the metabolism of the widely used proton pump inhibitor drugs such as omeprazole, pantoprazole, or lasoprazole is mainly performed by the *CYP2C19* isoform. The main route for the metabolism of omeprazole is converting the substance to hydroxyl and 5-O-demethyl metabolites by *CYP2C19* isoform, and further to 5-hydroxyomeprazole sulfone by *CYP3A4*[130, 131]. However, omeprazole can also be converted by *CYP3A4* to omeprazole sulfone and then to 5-hydroxyomeprazole sulfone by *CYP2C9*[132, 133]. The biotransformation of blood thinner drugs like S-warfarin is performed by the isoform *CYP2C9* to 7-hydroxywarfarin [134, 135]. On the other hand, its enantiomer R-warfarin is metabolized mainly by *CYP1A2* to 6-and 8-hydroxywarfarin and by *CYP3A4* to 10-hydroxywarfarin, and by carbonyl reductases to diastereoisomeric alcohols [136]. The biotransformation reactions of ß-blocker drugs like bufuralol or metoprolol are catalyzed in the liver by *CYP2D6* isoforms in their majority [137]. Concretely, metoprolol is transformed to α-hydroxymetoprolol and O-demethylmetoprolol primarily, and minorly by CYP3A4 to O-demethylmetoprolol [138]. The most abundantly expressed CYP in liver is the *CYP3A4* isoform, and the CYP3A subfamily metabolizes nearly 30% of all the used drugs from all therapeutic categories [139-141]. For instance, the benzodiazepine drug midazolam is transformed to 1-OH-midazolam and 1-OH-midazolam-glucuronide by *CYP3A4* and to a minor extent by *CYP3A5* [142]. These two isoforms share 85% amino acid sequence similarity, therefore sharing high substrate selectivity [143].

The substrates of the different isoform of the CYP450 enzymes are extensively used to phenotype the metabolic capacity of the liver in both *in vivo* and *in vitro* studies [123, 144], which will be further elaborated later in the introduction of this thesis.

Inter-individual variations in CYP450 enzymes: Single-nucleotide polymorphisms (SNPs)

The human interindividual variability described regarding the drug metabolism and response presents a challenge in clinical practice because it affects the biotransformation efficiency and rate and the response toward pharmacological treatments and therapeutic effects [145-147]. Overlapping substrate specificity, the presence of genetic single nucleotide polymorphisms (SNPs) [148] , and variations in the distribution among different ethnicities [149, 150] are responsible for this divergence [151]. High polymorphism rates have been found among the CYP450 genes, especially *CYP2B6, CYP2C9, CYP2C19,* and *CYP2D6* with at least 70 different allelic variants described [152] and for *CYP2C19*, which have been discovered by the phenotypic pharmacokinetic and pharmacodynamic differences they elicit in drug-treated individuals [148]. These influence the individual capacity to perform biotransformation reactions of certain compounds, dividing the population into: i) "poor metabolizers" (PM), carriers of null alleles with complete lack of function; ii) "extensive metabolizers" (EM) which refers to the most commonly found phenotype among the population; iii) "intermediate metabolizers" (IM), who carry one functionally deficient allele and one functional, leading to impaired drug oxidation capacity; and iv) "ultrarapid metabolizers" (UM) originating from variants leading to gain of function [117, 152-155].

1.3.2 Phase II biotransformation reactions

Hepatic phase II metabolism of xenobiotics comprises conjugation reactions catalyzed by transferase enzymes that serve for the solubilization and excretion facilitation of the products from phase I metabolism [114]. These reactions include glucuronidation, sulfation, methylation, acetylation, glutathione conjugation, and aminoacidic conjugation [156, 157]. Families of transferase enzymes carrying out phase II metabolism are UDP-glucuronidases (UGTs), sulfotransferases (SULTs), N-acetyltransferases (NATs), and glutathione S-transferases (GSTs), amino acid conjugation enzymes and various methyltransferases (thiopurine S-methyltransferase and catechol O-methyltransferases [156, 158, 159]. Some of the phase II metabolizing enzymes have also been shown to be related to the stress response upon drug treatment.

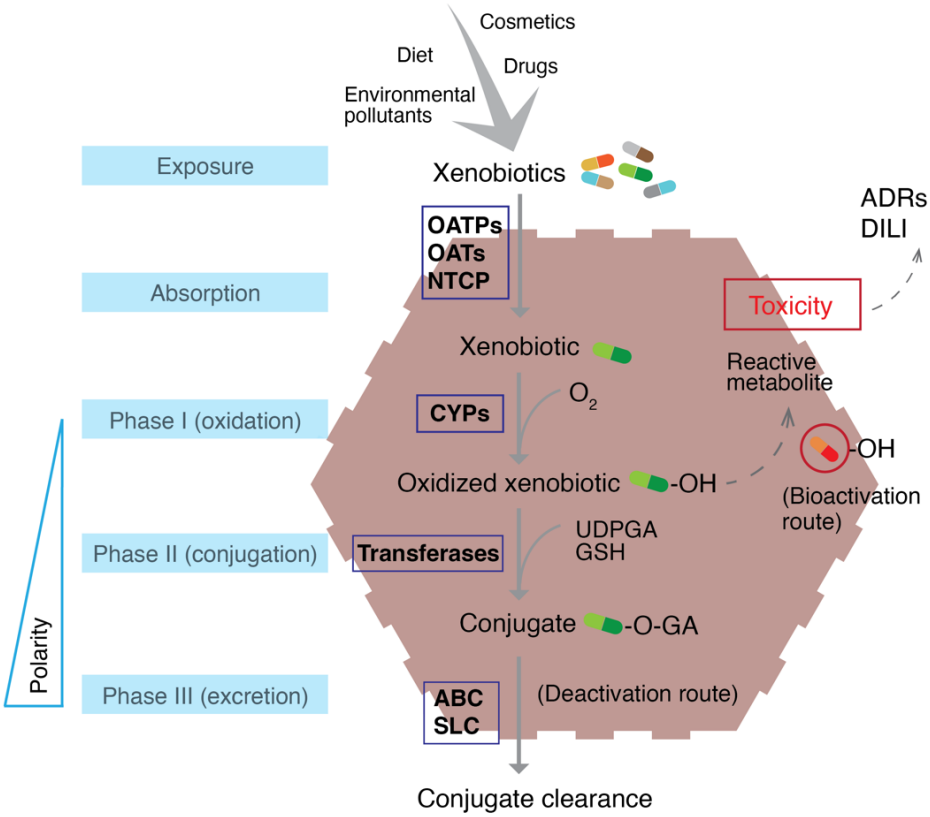


Figure 2. Schematic representation of the xenobiotic metabolism in hepatocytes. Xenobiotics are transported into the cells through organic anion transporters (OATs) and organic anion transporting polypeptides (OATPs) and sodium taurocholate co-transporting polypeptide (NTCP). *During phase I, CYP450s enzymes catalyze biotransformation reactions such as hydroxylation as exemplified here, in the presence of oxygen. Phase II metabolism is carried out by transferases including glutathione S-transferases (GST), methyltransferases, glycine N-acyltransferase (GLYAT), N-acetyltransferases (NAT), sulfotransferases (SULT), UDP-glucuronosyltransferases (UGT) by incorporating uridine diphosphate-glucuronic acid (UDPGA) and using glutathione (GSH). Phase III refers to the conjugate excretion through ATP binding cassette (ABCs) (e.g., multidrug resistance protein family)) and solute carrier transporters (SCLCs) transporters. GA: glucuronic acid; GSH: glutathione; OH: hydroxyl; PAPS: phosphoadenosine-phosphosulfate. Image adapted from Esteves et al. 2021[160].*

1.3.3 Phase III drug metabolism

During this phase, the conjugated xenobiotics are transported out of the cell across the cellular membrane in an ATP-consuming protein [114]. Four different classes of membrane-bound protein transporters exist: ion channels; transporters; aquaporins; and ATP-dependent pumps. ATP-binding cassette (ABC) transporters are an example of ATP-dependent pumps. The major proteins exporting drugs outside of the cell are ABC transporters that are located in the canalicular membrane of the liver. For instance, MRP2 and BCRP, export conjugated compounds into the bile. The main members of the ABC family are ABCA, ABCB and ABCC, and their inhibition leads to an increase in the concentration of their substrates in the bloodstream and decreases their biliary excretion, prolonging their stay in the body [161].

1.4 The aging liver and the aging liver in disease

The coadministration of drugs is highly common in the geriatric clinical practice in order to treat age-related comorbidities such as hypertension, cardiovascular disease, dyslipidemia [162], type 2 diabetes, or vascular diseases [163, 164]. Hence, the elderly population is in a more susceptible position to developing ADRs, subsequent hospitalization, and acute liver failure [165, 166]. In fact, an incidence of ADRs in the elderly twice as high as in younger patients has been registered [167]. In addition, polypharmacy in generally associated with adverse drug outcomes, including DILI and DDIs [168].

In general, aging is marked by a decline in the function of organs and organisms. Among the described hallmarks of aging, there have been stem cell exhaustion, cellular senescence, mitochondrial dysfunction, deregulated nutrient sensing, loss of proteostasis, genomic instability, telomere attrition and epigenetic alterations [169]. The aging process in the liver is governed by transcriptomic and epigenomic alterations that contribute to the dysregulation of mitochondrial function and nutrient sensing pathways and energy homeostasis, resulting in cellular senescence and low-grade inflammation [170]. In addition, aging is associated with the metabolic syndrome, defined by the World Health Organization (WHO) as a pathologic condition characterized by abdominal obesity, hypertension, insulin resistance and hyperlipidemia, as well as to increased chronic inflammation [170-172]. In fact, cellular senescence has been suggested to drive age-dependent hepatic steatosis [173, 174]. A general decline in tissue functionality and performance capacity has been described, such as a lower blood flow with age [175]. Alongside the increase in age, there is an increase in hepatic lipid accumulation and lipid metabolism in adipose tissue dysfunction [176], leading to impaired physiological defense against injuries the development of the metabolic syndrome and enhanced pathological pathways such as lipid accumulation [177, 178], developing into the phenotypic characteristics of NAFLD in the elderly population, among which a higher incidence has been registered [179, 180]. Furthermore, the inflammation and ER stress related to fat accumulation and to aging promotes the expression of heat shock proteins, and their accumulation triggers the unfolded protein response (UPR) proteins [170, 181, 182]. The accumulation of misfolded proteins and increase in lipogenesis and lipotoxicity results in higher ROS levels, triggering oxidative stress and anti-inflammatory stress response [183], leading to an overall higher age-related liver diseases predisposition, such as steatosis and NAFLD [170].

1.4.1 Aging and variability

Aging has been associated with an increase in transcriptional cellular variability in several tissues and organs [184]. Martinez-Jimenez *et al.* demonstrated that aging increased the cell-to-cell transcriptional variability and uncoordinated responses in CD4T+ immune T cells [185]. This has also been described in lung [16], heart [186], and human pancreas [187]. An extensive atlas of the mouse tissues during aging (*Tabula muris senis* or 'Mouse Aging Cell Atlas') has been established by the *Tabula Muris Consortium*, where the authors described a decrease in the number of genes per cell detected in older animals across most murine tissues, including liver [188]. Furthermore, changes in the cellular composition during aging have been described in the liver, resulting in an increase in the inflammation-related macrophages that drive cellular senescence [188, 189].

In addition, evidence indicates that age-related changes in the liver play a key role in the susceptibility to develop NAFLD [173]. In fact, a higher incidence of NAFLD in the elderly has been reported repeatedly in the literature [179, 190]. Changes observed at the transcriptomics level and functional responses are modulated by the chromatin configuration at the epigenomic level. Moreover, increased chromatin accessibility and lower polymerase pausing activity were described as associated to aging [191]. It has been postulated that cell-to-cell epigenomic variability or noise is one of the mechanisms that might lead to higher transcriptional noise with age [184, 192]. For instance, a recent multiomics study of the liver by Nikopolou *et al.* demonstrated that hepatocytes showed a clear separation between young and old cells driven by their chromatin accessibility profile assessed through scATAC-seq, however not in their transcriptomic profile. Nevertheless, they appointed that age still represents a relevant factor for explaining transcriptional variability between cells [193].

The impact of aging on the changes in transcriptional responses to exposure to several drugs simultaneously causing a drug-related metabolic challenge has not been explored in individual PHHs.

1.5 Liver in disease: fat accumulation and NAFLD

The intracellular accumulation of fat in the hepatocytes is known as hepatic steatosis, when 5% or more of the hepatocytes present triglyceride accumulation exceeding 5% by weight without an excessive alcohol consumption (maximum of <20 g/day for females and <30 g/day for males) [194-196]. Steatosis is a hallmark of NAFLD, the most common liver disease, affecting around 25% of the population worldwide [197]. In approximately 10 to 20% of the patients, it can develop into

non-alcoholic-Steatohepatitis (NASH), which can progress to cirrhosis and eventually resulting in hepatocellular carcinoma (HCC) [198].

The occurrence of NAFLD is tightly associated to insulin resistance and glucose intolerance, which in general constitute the characteristics of the metabolic syndrome and Diabetes Mellitus, as well as cardiovascular-related diseases like coronary heart disease (CHD) [199]. In addition, this prevalence increases to 90% in obese patients [200] and is also correlated with severe comorbidities like dyslipidemia (high plasma TG and/or low plasma HDL-cholesterol concentrations), and hypertension [201].

Fat accumulation is triggered in the liver by a decompensation in the lipid metabolism, such as increased uptake of circulating fatty acids, increased *de novo* lipogenesis, decreased hepatic β-oxidation and decreased lipid export [202] (**Figure 3**). This occurs with a concomitant increase in short-chain and monounsaturated fatty acids [203] and decreased levels of triglycerides containing polyunsaturated fatty acids (PUFAs) such as ω-3 and ω-6 fatty acids [203]. The phenotype in NAFLD is reversible through lifestyle and dietary intervention, including weight loss and exercise, to manage associated conditions such as obesity, diabetes, and hyperlipidemia [201]. However, there are currently no approved pharmacotherapies nor known biomarkers to assess the progression of the disease or the regression of more severe disease statuses, such as NASH [204]. Furthermore, the accumulation of fat in the liver is associated with lipotoxicity, endoplasmic reticulum (ER) stress, inflammation and chemokine production, which are determinants for the progression of NAFLD into NASH [205]. ER stress also acts promoting hepatocyte cell death, systemic inflammation and insulin resistance by activating *Nrf2*, *JNK* and *NF-κB*, *CREBH*, and *CHOP*, which actively contribute to inflammatory processes and cell death, promoting disease progression [205]. This leads to the increase in oxidative stress and reactive oxygen species (ROS) production, contributing to the NAFLD inflammatory phenotype and to the progression to NASH [206](Figure 2). For instance, an increase in the transcription of *CYP2E1* has been recognized as a marker of the disease, as it leads to an increase of ROS and superoxide anion radicals' production, promoting the progression of NAFLD to NASH [207-209].

To date, studies at single-cell resolution of the effect of fat accumulation as occurring in chronic diseases such as NAFLD on the transcriptional responses of individual hepatocytes are scarce. These have been shown to increase the cellular heterogeneity in the liver, demonstrated in both mouse [106, 210] and human [2, 211, 212] liver. A study published by Su *et al.* in murine liver revealed heterogeneous transcriptional profiles among liver macrophages and seven subpopulations of hepatocytes along NAFLD progression upon high-fat diet feeding [210]. They also identified a subset of zone-specific genes activated upon high fat diet feeding. In humans, Xiong *et al.* identified a

NASH-associated type of macrophages (NAMs) in mouse and human chronic disease [211]. Addressing the progression of NAFLD, Wang et al. identified fibrosis-associated genes that may serve as druggable targets against NASH [213]. Ramachandran *et al*. resolved the human cirrhotic liver niche, focusing on the non-parenchymal cell compartment [8], and Aizarani *et al.* discovered perturbed cellular phenotypes and gene signatures in HCC [2]. At the epigenetic level, DNA methylation patterns as well as histone modifications and microRNA expression have been associated to the development of metabolic-induced fatty liver and other metabolic diseases such as obesity or diabetes [214]. Moreover, at the epigenomic level, using a rat model, distinct *in vivo* transcription factor modulation and core genes for NAFLD have been recently discovered [215, 216]. In humans, scATAC-seq analysis of healthy, steatotic or fibrosis NASH livers revealed differentially accessible regions profiles with stage-specific DNA regulatory elements allowing NAFLD subtypes [217].

All in all, heterogeneity in the intracellular lipid accumulation pattern has been histologically described in terms of the number and size of the lipid droplets in hepatocytes [218], but whether chronic lipid accumulation triggers harmonized and coordinated transcriptional responses, or increases cellular variability among human hepatocytes remains thus far unclear.
In this thesis, scRNA-seq is used to assess the metabolic profile of individual PHHs *in vitro* upon chronic lipid accumulation, their transcriptional responses to a drug cocktail, and the exposure to these two metabolic challenges simultaneously. Moreover, scATAC-seq is applied to address how chronic lipid accumulation impacts the genome-wide chromatin accessibility configuration, which is largely unexplored at single-cell resolution.

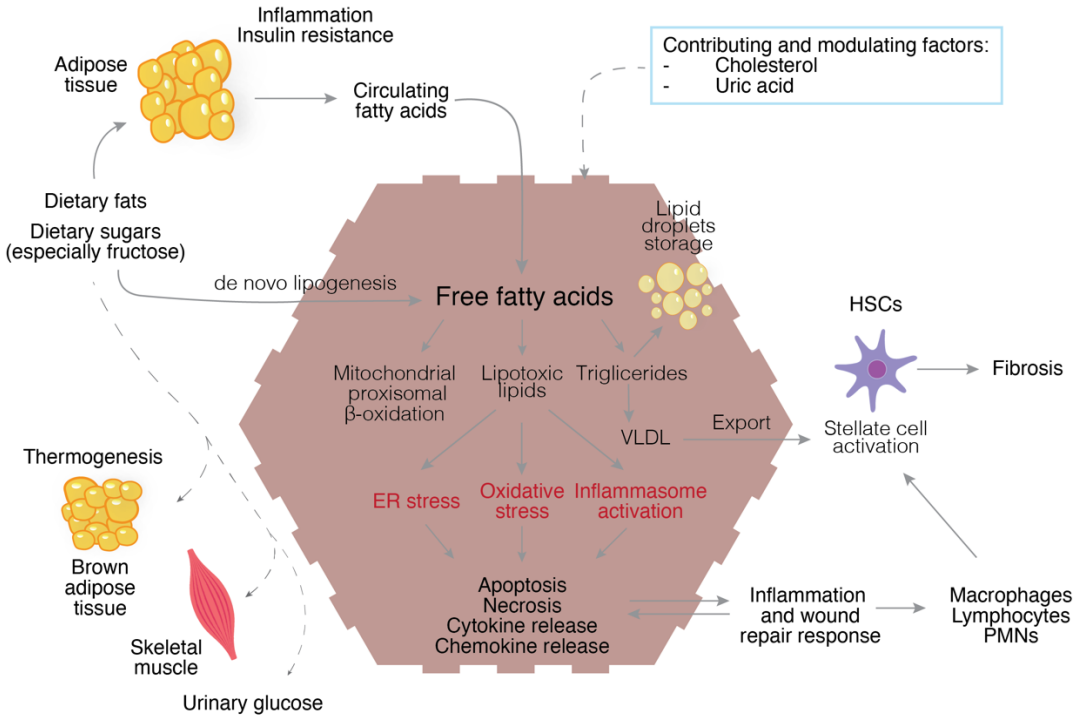


Figure 3. Graphical summary of the intracellular lipid metabolism in hepatocytes contributing to the development of the phenotype observed in hepatic steatosis, NAFLD and NASH. *Fatty acids play a key role in the pathogenesis. Free fatty acids originate from lipolysis of triglycerides in adipose tissue and transported in blood. De novo lipogenesis is a major contributor to the generation of free fatty acids since excess carbohydrates are converted to fatty acids. Fatty acids can then be oxidated in mitochondria and re-esterified to form triglycerides. These can be exported to the blood stream as VLDL or stored in lipid droplets intracellularly. The triglycerides stored in lipid droplets can then undergo regulated lipolysis to release free fatty acids again. If the B-oxidation of fatty acids is saturated, these can contribute to the formation of lipotoxic species that lead to ER stress, oxidative stress and inflammasome activation. These processes are responsible for the NASH phenotype, occurring with inflammation and wound repair responses activation, stellate cell activation and resulting in fibrosis. Image adapted from Freidman et al. 2018 [219].*

1.5.1 Drug metabolism alterations in NAFLD

Hepatic steatosis alters the physiological drug-related metabolic capacity of the liver [220]. The drug-related metabolic capacity of the liver has been reported to be diminished in obese and NAFLD patients and in animal models. This reduced capacity for drug metabolism leads to a higher concentration of the compounds in the bloodstream and leads to a higher risk of acute toxicity and worsening symptoms induced by commonly taken drugs, such as acetaminophen, losartan and anesthetics [220, 221]. Significant dysregulations in gene expression, protein levels, and enzymatic activity of multiple CYP450 enzymes have been described in several studies in liver microsomes and other *in vitro* models of NAFLD [207, 222, 223]. For instance, the gene expression, enzymatic activity, and protein levels of the major enzyme *CYP3A4* have been shown to be impaired and significantly reduced in human livers [207 , 224-226]. Phase II enzymes have also been shown to be impaired in fatty livers, for instance showing a reduced enzymatic activity of *SULT1A2* [227] and *GST* [223]. However, there is to date an overall lack of knowledge on the inducibility of these enzymes by drugs in the NAFLD liver

at single-cell resolution [228]. The dissection of the transcriptional responses of individual hepatocytes in response to the exposure to several drugs at the same time when the cells are loaded with intracellular lipids as in chronic diseases such as NAFLD has to date not been performed.

The present thesis aims for the study of the metabolic capacity of individual primary human hepatocytes exposed to chronic fat accumulation, and to the simultaneous challenge of a drug cocktail and chronic intracellular lipid accumulation. In addition, the chromatin accessibility landscape governing gene expression upon fat accumulation, a drug-related challenge by exposure to several drugs simultaneously, or both concomitantly is assessed in individual PHHs *in vitro*. The exploration of the gene regulatory landscape allows for the identification of differential chromatin openness patterns in response to intrinsic factors affecting cellular heterogeneity, such as chronic intracellular lipid accumulation, extrinsic factors such as exposure to a drug cocktail or both simultaneously. These open chromatin configurations determine the gene expression profiles observed downstream in transcriptomic analyses. In order to analyze chromatin accessibility, the nuclei isolation procedure for scATAC-seq on PHHs *in vitro* cultured is optimized for the application of high throughput, droplet-based approach (10X Genomics).

1.6 *In vitro* models of liver and steatosis

1.6.1 Hepatic immortalized cell lines

Several human hepatoma cell lines have been carefully studied and analyzed in order to serve as models of liver cultured in monolayer, among which HepG2 or HepaRG are the most commonly used [229-233]. The HepG2 cell line presents several limitations, such as the lack of nuclear receptors such as the nuclear receptors governing of drug-metabolizing enzyme's activity, namely CAR or PXR [234-236]. In addition, they lack the expression and induction capacity of phase I metabolizing enzymes. Their genome dysregulation leads to genotype instability with increased passaging, making them an unsuitable model for induction studies to predict drug metabolism [237-242]. A possible solution is a transduction with adenoviral vectors encoding CYP450 genes, however, its inconsistent efficacy and technically demanding experimental conditions make this a difficult strategy for simultaneous CYP450 assessment or drug-metabolism-related assessments [243]. HepG2 cells still present a suitable model for chronic intracellular lipid accumulation studies, since they express the necessary enzymes involved in lipid metabolic pathways [244].

HepaRG cells have been established as the closest model to primary human hepatocytes (PHH), since they maintain expression levels of CYPs and hepatic key nuclear receptors, as well as membrane

transporters and phase II enzymes [231, 233, 238]. Induction studies using probe substrates of the main drug-metabolizing phase I enzymes CYP1A1, CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, and CYP3A4 showed induction of their mRNA level upon exposure, making them a valuable *in vitro* model for human drug metabolism studies [245-248]. An added advantage without presenting the inter-donor variability and functional instability of PHHs [240]. In addition, they resemble the morphology of primary hepatic cells by maintaining the trabecular structure. However, these cells, as occurring also with HepG2 cells, were not as indicative as PHH when predicting drug-induced hepatotoxicity [242]. Therefore, they also do not constitute a complete model for the investigation of hepatotoxic responses.

1.6.2 Primary Human Hepatocytes (PHHs)

Liver biopsies can be used as a source to obtain primary cells, in particular, PHHs are considered the gold standard model for *in vitro* drug testing and cytotoxicity studies [238]. Primary hepatocytes are classically isolated from fresh liver tissue through the two-step collagenase perfusion method [249, 250]. They are considered the gold standard because they retain the characteristics and functionality of liver tissue, and importantly, high phase I and phase II enzyme expression levels, representing the closest model to *in vivo* and a suitable model to be used for enzyme induction and inhibition studies [251, 252]. However, the inter-donor variability and their scarce availability together with the difficulties they present for long-term culture maintenance make them challenging to work with [240]. Another of their limitations is that along the incubation time, in both 2D monolayer culture, 3D or sandwich culture modalities, the hepatocytes undergo a de-differentiation process whereby they lose their hepatic phenotypic characteristics including morphology, structure, polarity and gene expression profile [242, 253-255]. As the isolated cells have lost their microenvironment structure, cell interactions, membrane structures, and cell functionality are decreased over time [237, 252]. In addition, metabolic zonation patterns are lost due to the limitations of conventional culture conditions, which do not provide the gradient of nutrients needed to maintain zonation-like phenotypes [256]. For instance, the key hepatocyte marker protein albumin production, or CYP450 expression quickly decline in the first 24 to 48 hours of incubation, leading to the cells losing their differentiated and mature hepatocyte signature [257]. Nevertheless, in the liver and in PHHs, the mRNA and protein levels of CYPs show a highly significant correlation for *CYP1A1, CYP1A2, CYP3A4, CYP2D6* and

CYP2B6 [123, 258], indicating that their mRNA levels can be used as an accurate estimate of their activity [259].

Hepatic steatosis in cultured human cells can be mimicked and induced *in vitro* to recapitulate and closely model the intracellular lipid accumulation phenotype occurring in human NAFLD [260, 261]. The induction of steatosis can be achieved by incubating the cells with free fatty acids (FFA) in the culture media, which is the most extensively studied and practiced strategy [260, 262-266]. *In vitro*, cells incubated with free fatty acids store triglycerides in organelles named lipid droplets [267, 268]. The incubation of cells with different ratios of the mixture of saturated and unsaturated FFAs triggers different responses in the cells. Additionally, there is heterogeneity among the way cells accumulate lipids in regards of the number and size of lipid droplets [218]. Whether all cells respond to lipid accumulation in a coordinated fashion still remains unclear. For instance, a high concentration of saturated fatty acids exerts cellular lipotoxicity [269]. The cytotoxic effects of saturated fatty acids can be buffered by adding unsaturated fatty acids in the culture media, therefore the ratio of the mixture of these two determines the effect of lipid accumulation in the cells. Excessive intracellular lipid accumulation can lead to lipotoxicity, mediated by increased ER stress, and oxidative stress and culminating in apoptosis [270].

1.7 Drug-induced liver injury and toxicity – bulk vs. single-cell approaches

1.7.1 Drug-drug interactions and phenotyping cocktails

Classically, in preclinical early phases of drug discovery and development PHHs are used as the gold standard model for the assessment of drug efficacy and safety [238]. Since multiple drug therapies are common in clinical practice, the assessment of the treatment's safety is crucial. During these phases, the assessment of drug efficacy, potential drug-drug interactions (DDIs), hepatotoxicity and pharmacokinetic and -dynamics are performed in bulk analyses and measurements, considering that hepatocytes are a seemingly homogeneous population of cells [237, 271]. However, whether the molecular phenotype and functional responses are shared among all hepatocytes *in vitro* is yet to be elucidated.

Drug phenotyping cocktails have been developed and applied for the assessment of the metabolic capacity of the liver, appointing for the known as the "cocktail approach" [272-274]. This strategy was adopted to overcome the disadvantages that assessing each CYP450 isoform individually poses, such as the elevated number of samples, time and labor consuming and cost infectivity [275]. The

concomitant administration of several probe substrates provides the advantage of assessing the drug interactions and CYP450 activity more efficiently [276, 277].

In vitro, changes in the mRNA levels upon incubation with the substrates can be used to phenotype and measure the metabolic functionality of the liver. Cytochrome induction and inhibition determine the safety and further development of drug candidates [278]. A myriad of phenotyping drug cocktails has become available, composed of different CYP inducers in a variety of doses and concentrations. Some examples are the firstly developed Pittsburg cocktail, followed by the Cooperstown cocktail [279, 280], the Karolinska cocktail [281], and the Inje cocktail [282]. Others such as the Geneva cocktail [283], the Basel cocktail [284], and the Sanofi-Aventis cocktail [285] have also been developed targeting other CYP isoforms and tackling the limitations of previously published ones. The latter, the Sanofi-Aventis cocktail, is composed of five probes that are inducer substrates of the main CYP450 enzyme isoforms, responsible for the metabolism of 70-80% of the clinically used drugs in the market: caffeine (*CYP1A2*), omeprazole (*CYP2C19*), S-warfarin (*CYP2C9*), metoprolol (*CYP2D6*) and midazolam (*CYP3A4*) [125, 285-287].

In humans, the isoform CYP1A2 is responsible for 90% of the caffeine metabolism, catalyzing the N-3 demethylation of caffeine to paraxanthine. theophylline (12%) and theobromine (4%) [288-290]. In addition, CYP2E1 contributes in a lesser extent to the formation of theophylline and theobromine [290, 291]. It has also been reported that caffeine has binding sites in recombinant CYP3A4, modulating its enzymatic activity [292].

The proton pump inhibitor drug omeprazole is primarily metabolized by CYP2C19 generating 5-hydroxiomeprazole and secondarily by CYP3A4, producing omeprazole sulfone [293]. It is known that omeprazole, which is mainly metabolized by CYP2C19, and its metabolites are time-dependent inhibitors of CYP2C19 and CYP3A4 in a reversible manner [294]. This means that omeprazole inhibits its own metabolism, leading to non-linear pharmacokinetics [153]. Moreover, omeprazole has been reported to induce CYP1A2 in a time-dependent manner, depending on its clearance rate by CYP2C19 [295].

The anticoagulant drug warfarin is composed of a racemic mixture of R- and S-warfarin. These two are metabolized differentially by human CYP450 enzymes. The racemic isomer S-warfarin, is efficiently metabolized by the isoform CYP2C9, leading to the formation of 7-hydroxywarfarin [296], whereas CYP2C19 and CYP3A contribute in a minor manner to other metabolic pathways [134, 297, 298].

The isoform CYP2D6 catalyzes the metabolism of beta-blocker drug metoprolol (70-80%) [299] to produce O-demethylmetoprolol and α-hydroxymetoprolol, to which CYP3A4, CYP2B6, and CYP2C9

contribute to a lesser extent *in vivo* [300]. Further oxidation of O-demethylmetoprolol forms metoprolol phenylacetate [300]. Metoprolol and its metabolites have been reported to inhibit CYP2D6 and CYP3A4, without affecting the CYP2D6-mediated metabolism of midazolam [301].

The benzodiazepine drug midazolam is metabolized in its majority by CYP3A4 to form the principal metabolites 1'-OH and 4-OH midazolam [302]. This drug is considered to have high specificity as no other CYP450 isoforms have been described to intervene significantly in the metabolism of midazolam [302, 303].

Taken together, the use of substrates of different CYP450 isoforms entails the possible inhibition of other CYPs, leading to potential undesirable drug-drug interactions when administering several drugs simultaneously. Therefore, the selectivity and the specificity of the cocktail components and the doses or concentration are of crucial importance to avoid drug interactions and biased readouts [277]. The Sanofi-Aventis cocktail has been demonstrated *in vivo* to not interfere with the clearance of the other probe drugs [285] and to show high safety and specificity in both *in vitro* [304] and *in vivo* [285, 287, 305] experiments. In this thesis, the Sanofi-Aventis cocktail is used to phenotype the drug-related metabolic ability of individual primary human hepatocytes *in vitro*, exposing them to a metabolic challenge.

1.8 Hypotheses and aims of this thesis

The hypotheses stated for the research projects constituting the main focus of this thesis are that intrinsic and extrinsic factors may alter the gene expression profile and the chromatin accessibility landscape of single primary human hepatocytes *in vitro*. These factors include chronic intracellular lipid accumulation, aging, exposure to drugs or both simultaneously *in vitro* (**Figure 4**). Thereby, individual hepatocytes may show differential transcriptomic responses toward a metabolic challenge such as exposure to a five-drug cocktail. Moreover, intracellular fat accumulation as occurring in NAFLD may impact the metabolic profile of single hepatocytes, and their capacity to metabolize a cocktail of drugs simultaneously. Presumably, changes at the transcriptomic level might be reflected at the epigenomic level, leading to heterogeneous transcriptomic responses and epigenomic changes upon exposure to these factors.

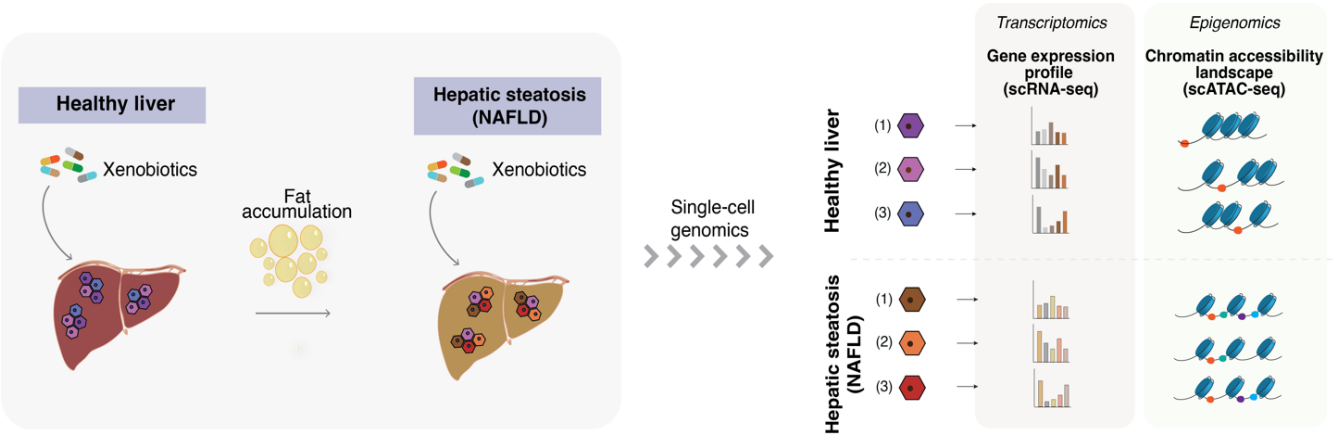


Figure 4. Main hypothesis stated in the present thesis. *Chronic intracellular lipid accumulation may impact the xenobiotic-related metabolic capacity of single human hepatocytes. Single-cell genomics technologies are applied to study the gene expression profile (transcriptomics) and the chromatin accessibility landscape (epigenomics) of individual hepatocytes in vitro. Presumably, heterogeneity in the functional drug-related metabolic profile of single cells can be found in healthy liver and perturbed in NAFLD.*

The development of single-cell technologies has revealed previously concealed characteristics of tissues and individual cells. In this thesis, the aims are:

a) To characterize rare hypothalamic oxytocin-expressing neurons in the mouse brain at the cellular level. The snRNA-seq2 methodology is applied to characterize the nuclear transcriptome and identify the transcriptomic responses associated with an obesogenic high-fat/high-sugar diet (Chapter I).

b) The development of a plate-based methodology for the assessment of the epigenetic landscape of complex tissue features such as liver polyploidy (Chapter II).

c) The optimization of a high-throughput droplet-based protocol for the exploration of the chromatin accessibility at single-cell resolution (scATAC-seq) of primary human hepatocytes *in vitro*, in four experimental conditions (Chapter III).

d) The deep characterization of the transcriptional response of individual primary human hepatocytes *in vitro* in response to exposure to a drug cocktail, chronic intracellular lipid accumulation, or exposure to drugs to lipid-laden hepatocytes using a high throughput scRNA-seq approach (Chapter IV).

2. Material and Methods

This section includes the methodology used to address the questions and hypotheses stated for this project. The outline of the procedures and methods used for this thesis are divided into the two single-cell genomics approaches used: plate- and droplet-based approaches. Firstly, I proceed to explain the methodology used for the single-nucleus RNA-sequencing of hypothalamic neurons. Secondly, I detail the protocol followed for single-nucleus ATAC-seq in plates.

2.1 Low throughput plate-based methodologies on mouse tissue: snRNA-seq and scATAC-seq

2.1.1 Single-nucleus RNA-seq2 applied to mouse hypothalamic neurons

This project was a collaboration with Dr. Tim Gruber and the Dr. García-Cáceres lab, the Astrocyte-Neuron Networks group at the Institute for Diabetes and Obesity at the Helmholtz Zentrum München, published in Gruber *et al.* 2023 [306].

The overview of the experimental methodology to assess the cellular heterogeneity among murine hypothalamic neurons is depicted in **Figure 5**. Nuclei from flash-frozen hypothalamic tissue from mice fed either a chow or an obesogenic high-fat/high-sugar (HFHS) diet were isolated. Next, oxytocin and GFP-positive neurons were FACS sorted into 384-well plates, half a plate per treatment. Subsequently, with the aid of a liquid handling robot, the snRNA-seq2 methodology was minutely followed.

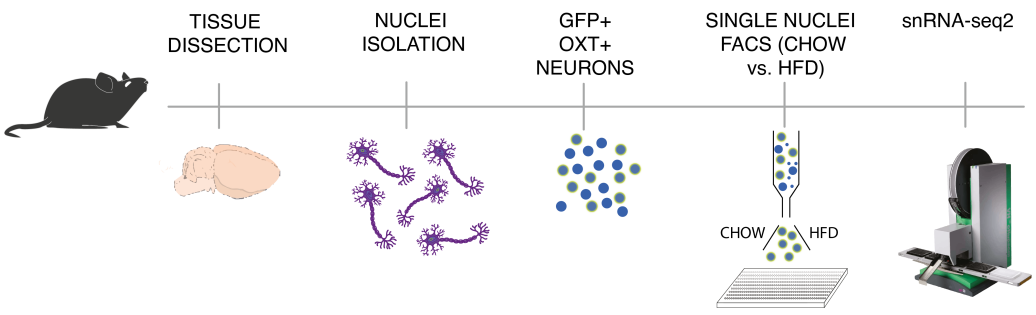


Figure 5. Schematic illustration of the experimental design for the analysis of hypothalamic Oxytocin-positive, GFP-positive neurons using the snRNA-seq2 method.

Nuclei Isolation

The nuclei isolation was performed by Dr. Raian E. Contreras (Institute for Diabetes and Obesity, Helmholtz Zentrum München) [37].

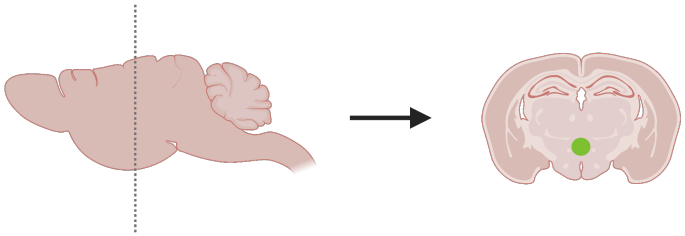


Figure 6. Schematic representation of the anatomical location of the hypothalamus in mouse brain. *Nuclei were isolated from flash-frozen hypothalamic tissue of genetically engineered (INTACT) mice.*

In brief, CAG-Sun1-sfGFP mice (INTACT mice) were crossed with OT-*ires*-Cre mice to generate heterozygous mice in Dr. García Cáceres lab [307]. Whole hypothalami (**Figure 6**) were individually processed to obtain single nuclei following a previously described protocol by Krishnaswami *et al.* 2016 [64] with minor modifications. In brief, frozen hypothalamic were transferred to a Dounce homogenizer containing 1 mL of freshly prepared ice-cold nuclei isolation buffer (0.25 M sucrose, 5 mM $MgCl_2$, 25 mM KCl,20 mM Tris pH 8.0, 0.4% IGEPAL 630, 1 mM DTT, 0.15 mM spermine, 0.5mM spermidine, 1x phosphatase and protease inhibitors, 0.4 units RNasin Plus RNase Inhibitor and 0.2 units SUPERase In RNase inhibitor (Life Technologies, AM2696)). To homogenize the tissue, 10 strokes with the loose pestle ("A") were performed, followed by 5 min incubation on ice and 15 more strokes with the tight pestle ("B"). The homogenate was filtered through a 20 µm cell strainer, and centrifuged at 1000 rcf for 10 min at 4°C. The resulting nuclei pellet was resuspended in 450 µL of Staining Buffer containing PBS, 0.15mM spermine, 0.5mM spermidine, 0.4 units RNasin Plus RNase Inhibitor, 0.4% IGEPAL-630, 0.5% BSA supplemented with DAPI 1µg/µL, and incubated for 15 min on ice. Nuclei integrity was assessed in the DAPI channel under a Zeiss microscope (Axio Scope, Zeiss, Germany). Doublet discrimination and DAPI staining were used for appropriate gating of single nuclei and the signal on the 488 (FITC) channel of the IgG-isotype control determined the adequate gating of GFP^- and GFP^+ nuclei.

Flow cytometry – Sorting of single nuclei into 384-well plates

GFP^+ nuclei were sorted with a 85 µm nozzle into 384-well PCR plates (thin-walled, BioRad, HSP3901) prepared freshly with 940 nL of Lysis Buffer 1 (1 µL of 10X reaction buffer is diluted in 2,75 µL of water) (Takara kit SMART-Seq v4 Ultra Low input RNA) per well, aliquoted with the Mosquito HV (STP Labtech) liquid handling robot. The Reaction buffer was prepared following the manufacturer's instruction adding 1 µL of RNAse Inhibitor in 19 µL of 10X Lysis Buffer. A detailed description of the gating strategy defined to sort GFP+ nuclei can be found in Dr. Raian E. Contreras' doctoral dissertation [37].

Maximum sorting accuracy was ensured by a colorimetric assay with tetramethylbenzidine substrate (TMB, BioLegend, Ref. 421501) and 50 µg/mL of Horseradish Peroxidase (HRP, Life Technologies, Ref. 31490) [308].

In the plate layout, we sorted nuclei from Chow diet-fed animals in half of the 384 well plate and nuclei from HFD-fed animals in the remaining half. After sorting, every plate was firmly sealed (MicroAmp Thermo Seal lid, #AB0558), shortly vortexed for 10 s, centrifuged (4 °C, 2000 x *g* for 1 min), frozen on dry ice, and stored at −80 °C, until cDNA synthesis. A total of four 384-well plates were sorted for this study.

Single-nucleus RNA-seq2 (snRNA-seq2)

The single-nucleus-RNA-seq2 methodology was used to capture a high number of transcripts from frozen tissues, allowing for the generation of double-stranded full-length cDNA as described and detailed by Richter*, Deligiannis*, *et al.* 2021 [19]. In brief, the reaction volumes were miniaturized with the aid of the Mosquito HV robot. Per well, 2190 nL of Lysis Buffer 2 (LB2) were dispensed. The final volume of the mixture of Lysis Buffer (LB1 and LB2) was 3.125 µL, containing NP40 2% (Life Technologies 85124), Triton-X100 1%, 1/300,000 diluted ERCC RNA Spike-In-Mix (Life Technologies, 4456740), 3′ SMART-seq CDS Primer II A and RNAse-free water.

Every flash-frozen sorted plate was thawed directly on a – 20 °C chilled metallic holder while LB2 was added by the Mosquito HV robot. The plate was immediately sealed, vortexed for 20 s at 2000 rpm, centrifuged at 2000 x *g* for 30 s at 4 ºC and placed in a 72 °C for 6 min. ERCC spike-ins (Thermo Fischer Scientific, Ref. 4456740; Lot num 00892098) were 1:10 diluted, with RNAse-free water with 0.4 U/µL Recombinant RNase Inhibitor (Takara Clontech, Ref. 2313A) and a fresh dilution of 1 in 300,000 was prepared before the first strand synthesis.

Reverse transcription and Pre-PCR amplification steps were followed as described by the manufacturer with four times reduced volumes for all steps. The PCR program for the cDNA amplification was performed in a total of 21 cycles: 1 min at 95 °C, [20 s at 95 °C, 4 min at 58 °C, 6 min at 68 °C] × 5, [20 s at 95 °C, 30 s at 64 °C, 6 min at 68 °C] × 9, [30 s at 95 °C, 30 s at 64 °C, 7 min at 68 °C] × 7, 10 min at 72 °C. After cDNA synthesis, the yield was assessed in an Agilent Bioanalyzer with a High Sensitivity DNA kit.

Library preparation and sequencing

Sequencing libraries were prepared using the Illumina Nextera XT DNA Sample Preparation kit (Illumina, Ref. FC-131-1096) and the combination of 384 Combinatorial Dual Indexes (Illumina- Set A to D, Ref. FC-131-2001 to FC-131-2004). Using the Mosquito HV robot, the reaction volumes of

the Nextera XT chemistry were miniaturized, and the steps followed minutely as described by Richter et al. 2021 [19, 309].

In brief, 500 nL of the undiluted cDNA were transferred to a new 384 well-plate containing 1500 nL of Tagmentation Mix (TD and ATM reagents). All Nextera XT reagents (NT, NPM and i5/i7 indexes) were added stepwise to a final library volume of 5 µL per well. The final PCR amplification was performed through 12 cycles. Once the libraries were prepared, 500 nL from each well were pooled together into a tube (total volume of ~192 µL) to perform a final AMPure XP bead (Beckman Coulter, Ref. A63882) clean-up step. Two consecutive clean-ups with a ratio of sample to bead 0.9X led to library sizes between 200 and 1000 bp. The final libraries were assessed using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent). Prior to sequencing, the libraries were quantified using a Collibri library quantification kit (Thermo Fischer Scientific, Ref. A38524100) in a QuantStudio 6 Flex (Life Technologies) to ensure accuracy.

Each plate, with a total of 384 libraries, was pooled together in one final library. In total, 4 final libraries were sequenced using an Illumina NovaSeq 6000 NGS sequencer in an SP XP flowcell, in a paired-end 150 bases.

Read alignment, counting and filtering of the combined batches

The alignment of reads, filtering and normalization of the mouse hypothalamic neurons snRNA-seq2 data was performed by Dr. Viktorian Miok (Molecular Pharmacology laboratory at the Institute for Diabetes and Obesity in Helmholtz Zentrum München) and the produced figures specified in the Results section. The raw count matrix consisted of 1,536 single nuclei and 55,579 genes. Nuclei with >1000 and <4500 genes detected were kept whereas genes sequenced in less than 25 cells and with a read count below 250 were filtered out. After filtering, the final matrix was composed by 1,202 single nuclei and 13,867 genes, provided for the downstream analysis in the present thesis performed by myself. Marker gene identification was performed using the PanglaoDB list of marker genes from brain tissue [310], performed by myself.

2.1.3 Low throughput scATAC-seq using a plate-based approach

In order to complete the second aim of this thesis, I aimed to optimize the protocol for single-cell ATAC-seq using a plate-based approach.

Lysis plate preparation

One day prior to performing the experiment, lysis plates were prepared by aliquoting 2 µL 2X Lysis Buffer (100 mM Tris-HCl pH 7.4, pH 8.0, 100 mM NaCl (Sigma Aldrich, 59222C), 40 pg/mL Proteinase K (Ambion, AM2546) and 0.4% SDS (Life technologies, 15553027)) and 2 µL of 10 µM S5xx/N7xx Nextera Index Primer Mix (5 µM each) per well. The plates were then sealed and stored at -80ºC [60].

Nuclei isolation

Nuclei from frozen liver biopsies from wild-type C57BL/6 mice were isolated following minutely the snRNA-seq2 method [19]. The tissue was minced on a cold petri dish on ice in 1 mL of homogenization buffer containing 250 mM Sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris buffer pH 8.0, 1% DTT, one tablet protease-inhibitor (cOmplete EDTA-free, Life Technologies, 1187358001), 0.1 % Triton-X and 0.2 % NP-40. Minced tissue suspension was transferred to a pre-chilled 2 mL douncer homogeneizer using wide orifice pipette tips and 5 strokes were performed with the loose pestle followed by 15 strokes with the tight pestle. Cells were filtered through a sterile 50 µM cell strainer (CellTrics, Symtex, #04-004-2327), using an additional mL of HB buffer to rinse and wash the douncer and passed through the filter. Cells were distributed into clean 1,5 mL Eppendorf tubes and centrifuged down for 8 min at 1000 rcf at 4ºC. After centrifugation, the cells were centrifuged in a iodixanol gradient to remove tissue debris. The cells were resuspended in 250 µL of HB buffer and 250 µL of 50% iodixanol (Optiprep, D1556, Sigma Aldrich Chemie) added and pipette-mixed carefully. The mixture was poured onto 500 µL of 29% Optiprep that were previously placed in an empty 1.5 mL Eppendorf tube. To obtain clean a nuclei pellet, the cells were centrifuged at 13,500 rcf for 20 min at 4 ºC.

Transposition and tagmentation reaction

The pellet was then resuspended in 50 µL of tagmentation mix containing 12.5 µL 4X THS-seq TB buffer (Illumina), 5 µL digitonin (Life Technologies, BN2006), 27.5 µL nuclease-free water and 5 µL Tn5 transposase (Illumina, Nextera kit, Illumina Cat No. FC-121-1030). Nuclei integrity was inspected under the microscope using trypan blue.

The reaction was then placed in the thermomixer at 37ºC, shaking at 8000 rpm for 30 min.

To stop the transposition reaction, 50 µL of TSB buffer (10 mM Tris-HCl, pH 8.0 and 20 mM EDTA, pH 8.0) were added and incubated on ice for 10 min. Thereafter, 200 µL of Nuclei Storage Buffer (166.5 mM sucrose, 5 mM $MgCl_2$ (Gibco Life Technologies, AM9530G), 10 mM Tris buffer pH 8.0

(Invitrogen 15568025) with Hoechst 33342 at a concentration of 10 µg/mL (Thermo Fischer Scientific) were added to the 100 µL mixture and transferred to a FACS sorting tube.

FACS Sorting into 384-well plates

Tagmented nuclei were sorted into 384-well plates using a 100 µm nozzle in a FACSAria™ II sorter (Becton Dickison) (Software version: FACSDiva Version 6.1.3). Firstly, the population containing all cells was selected, a subpopulation excluding doublets was selected. After sorting, the plates were centrifuged briefly and incubated at 65ºC for 30 min to release the tagmentase. Then, 4 µL of 10% Tween-20 (Sigma Aldrich, 11332465001), 2 µL RNAse-free water and 10 µL NEBNext High-Feidelity 2X PCR MasterMix (New England Biolabs, M0544L) were added per well. Library amplification was immediately performed in a thermocycler following: 72ºC for 5 min, 98ºC for 5 min, [98ºC 10 s, 63ºC 30 s, 72ºC 20s] X 18. All wells were then pooled together with other wells containing the same number of nuclei sorted in them. DNA was purified with a ZYMO DNA clean and concentrator kit (ZYMO RESEARCH, D4014) following the procedure minutely, and eluted in 10 µL of Elution Buffer (Qiagen, 19086). Library quality control for integrity, complexity and size distribution was performed in a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent).

2.2 High throughput single-cell genomics using a droplet-based approach: scRNA-seq and scATAC-seq.

To explore the metabolic capacity of single hepatocytes, a droplet-based approach was used providing advantages such as a higher throughput, or the ability to distinguish subpopulations within the same cell type over plate-based methods [28]. The conditions for the methodology and procedures utilized on PHHs for single-cell experiments were at first empirically optimized in two human hepatic cell lines: HepG2 and HepaRG cells. Once optimized, single-cell experiments on PHHs were performed using the 10X Genomics Chromium platform.

2.2.1 Optimization of free fatty acid incubation and the cocktail approach in hepatoma cell lines

a) Bulk experiments on HepG2 cells

HepG2 cells were used to optimize the free fatty acid media. The concentration of the fatty acid mixture as well as the time of incubation were assessed.

Cell culture

HepG2 cells were kindly provided by the Molecular Pharmacology laboratory at the Institute for Diabetes and Obesity in Helmholtz Zentrum München, initially purchased from ATCC (Manassas, VA, USA). The cells were incubated with DMEM medium with 4.5% glucose (Life technologies, 41965062), supplemented with 10% Fetal Bovine Serum (FBS) (Gibco, 10500064), 1% L-Glutamine (Gibco, 25030024) and 1% Penicillin-Streptomycin (Gibco, 15070063). The maintenance culture was incubated in T-75 vent-capped flasks (Corning, 354485) and passaged every 2 days or at 65-70% confluency, determined under brightfield microscopical observation.

Free fatty acid incubation

HepG2 cells were incubated with a mixture of the unsaturated oleic acid to the saturated palmitic acid in a 2:1 ratio at a concentration of either 100 µM, 150 µM or 200 µM for 72 h (**Table 1**). Both palmitic and oleic acid were purchased in powder from Sigma Aldrich (P0500 and O1008, respectively) and dissolved in ethanol to prepare a 50 mM stock solution [262]. In order to facilitate FFA uptake, pre-bounding of free fatty acids to 1% bovine serum albumin (BSA) (Sigma Aldrich, 10735078001) in a 1:5 molar ratio was performed by heating the mixture at 40ºC for 2 hours [262].

Table 1. Summary of the volumes and components of the free fatty acid incubation media.

Concentration of the 2:1 FFA mixture	Palmitic acid (µL)	Oleic acid (µL)	10% BSA (µL)	Maintenance med (µL)	Total volume (µL)
100 µM	16.5	33	500	49,451.5	50,000
150 µM	25	50	500	49,425	50,000
200 µM	66.5	133.5	500	49,300	50,000

Staining of intracellular lipids

Oil Red O staining of fatty acids was used to assess the intracellular accumulation of lipids in HepG2 cells, testing three different concentrations of the 2:1 ratio mixture of the unsaturated oleic acid to the saturated palmitic acid in a 2:1 ratio, mimicking plasma levels of FFA in human steatosis [260].

A sterile Oil Red O Stock Solution was prepared by dissolving 0.7 g of Oil Red O powder (Sigma Aldrich, O-O625) in 200 mL of 100% isopropanol (Sigma Aldrich, 59300-2.5L-M) by stirring overnight on a rocker and subsequently filtering it through a 0.2 µm filter (Millipore, SLGVV255F).

To prepare the Oil Red O Working Solution, 4 parts of water were mixed with 6 parts of Oil Red O Stock Solution, vortexed for 1 min and let sit for 20 min at room temperature.

HepG2 cells were seeded in collagen I-coated 6-well plates (Corning, 356400) and incubated for 24 h to allow the cells to attach to the plate surface. FFA were then added to the culture media in a concentration of either 100 µM, 150 µM, and 200 µM the cells were incubated for 72 h. After incubation, the culture media was aspirated, and wells were washed three times with 1 mL of sterile PBS. Cells were fixed by adding 2 mL of 4% PFA, immediately removing it adding another 2 mL, and incubating it for 1 h at room temperature on a rocker. After discarding the PFA, 1 mL of 60% isopropanol was added per well and left to evaporate and dry completely. Once dry, 2 mL of Oil Red O Working Solution were added without touching the walls of the well, the Oil Red O solution was removed carefully, and the wells were washed with 2 mL of room tempered water four times.

b) Bulk experiments on HepaRG cells

Fully differentiated NoSpin™ HepaRG™ cells were purchased from Biopredic International (Lonza) (Catalog number "NSHPRG"), thawed, and plated following the manufacturer's recommendations. In brief, cells were thawed and counted to plate approximately 800,000 cells per well in a collagen I-coated 12-well plate (WVR, 734-0166) in MHTAP medium. After 24 h, the medium was changed to Maintenance/Metabolism Medium (421). After 96 h, the cells were incubated with either DMSO, Cocktail, FFA+DMSO, or FFA+Cocktail for 48 additional hours.

RNA extraction

HepaRG and PHHs were pelleted at 500 rcf for 5 min and flash-frozen in liquid nitrogen for further RNA extraction. Thereafter they were stored at -80ºC for later processing. The cells were thawed in 500 µL of Tryzol™ slowly pipett-mixing. Then, 100 µL of Chloroform-Phenol-Isoamilic (Thermo Fischer Scientific) were added, the tubes shaken by hand for 15 sec and incubated 2 – 3 min at RT. Samples were then centrifuged at 12,000 rcf for 15 min at 4ºC. After that, a clear transparent phase of approximately 250 µL appeared in the upper part of the tube, which was collected and transferred to a clean tube and placed on ice. Then, an equal volume as the collected phase of 2-propanol (Sigma-Aldrich) was added, followed by flicking the tube and centrifuging down at 12,000 rcf for 20 min at 4ºC. The resulting supernatant was aspirated carefully and 500 µL of pre-chilled pure 80% Ethanol (Sigma) was added, and the samples were centrifuged at 12,000 rcf for 20 min for 4ºC. The supernatant was aspirated carefully, and the open tubes left to air dry under the fume hood for 10 – 20 min (do not

place on ice). Once dry, the pellet was resuspended in 15 µL of RNAse-free water. The RNA integrity and concentration were measured using a Nanodrop spectrophotometer (Thermo Fischer Scientific).

cDNA synthesis

Isolated RNA was used to generate complementary DNA (cDNA) by using Superscript II transcriptase enzyme (Thermo Fischer Scientific). The master mix was composed of the reagents and volumes shown in **Table 2**.

Table 2. Master mix for the reverse transcription reaction of 1 µg of total RNA to cDNA.

Volume (µL)	Reagent	Final concentration
3	DTT 0.1 M	18.4 mM
3.75	dNTPs	500 µM
6	5X PCR buffer (Promega)	1X
1	Superscript II (200 U/µL) (Promega)	200 U
0.625	RNAse inhibitor (40 U/µL) (Takara)	25 U
0.75	Primer oligo dT14 (A/G/C) 30VN 120 µM	3 µM
1.1	Rnase-free H_2O	-

A volume containing 1 µg of total extracted RNA was calculated per sample and the master mix added. Thereafter, the reaction was incubated for 90 min at 42ºC degrees. To stop the reaction, incubate 5 min at 95ºC. Thereafter, the cDNA was stored at -20ºC or used directly for RT-PCR.

qPCR

Quantitative real-time PCR was performed using the cDNA synthesized from isolated RNA. A 1:20 dilution of the obtained cDNA was used as input for the RT-PCR. For normalization, a commercially available 5 donor RNA pool was purchased and added as an extra sample in the RT-PCR reaction. The master mix was composed of 1 µL RNAse-free water, 5 µL SYBR green (Life technologies, 4364344), 0.5 µL of 20 µM forward primer and 0.5 µL 20 µM reverse primer for one sample. A total of 7 µL of master mix and 3 µL of 1:20 diluted cDNA were loaded per well of a 384-well microtiter plate in duplicates. The PCR program for the RT-PCR was performed as follows: 10 min at 95 °C, [15 s at 95 °C, 1 min at 60 °C, 20 s at 68 °C] × 42, [15 s at 95 °C, 1 min at 60 °C, 15 s at 95 °C]. The reaction was performed in a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher Scientific). The RT-PCR primer sequences used in the reactions are shown in **Table 3**.

Table 3. Forward and reverse primer sequences used in the qPCR performed on HepaRG cells.

Gene target	Forward primer	Reverse primer
PGBD	CGGAAGAAAACAGCCCAAAGA	TGAAGCCAGGAGGAAGCACAG
CYP2C9	TGTGGTCCTTGTGCTCTGTC	CAGAGTGAACACAGGGCCAT
CYP2C19	ACGGATTTGTGTGGGAGAGG	TGATAGAAGGGCGGGACAGA
CYP3A4	AAGTGGACCCAGAAACTGCA	ACTTACGGTGCCATCCCTTG

Primer efficiency calculations

The expression level was calculated using the fluorescence data (Ct values) measured in the qPCR using the methodology published by Ramarkes *et al.* 2003 [311]. Three values after the threshold value were taken for calculation of the correlation (R^2), the slope, and the intercept per replicate. Thereafter, all the resulting intercept values per replicate were averaged. The averaged value was then used to calculate the efficiency ($E = 10^{slope}$) and No ($10^{intercept}$). The formula applied to calculate the efficiency of the primers used to perform qPCR

$$N = N_0 \, x \, E^{Ct}$$

where $N0 = 10^{intercept}$ and $E = 10^{slope}$. To normalize each value by the housekeeping gene (*PGBD*), the averaged efficiency of all samples (10^{slope}) was raised to the power of the threshold cycle (Ct) of the reference gene – the threshold cycle (Ct) of the gene being calculated, which would be:
$E^{(CtReference - CtTarget)}$.

The melting curves for the four targets are shown in **Figure 7**:

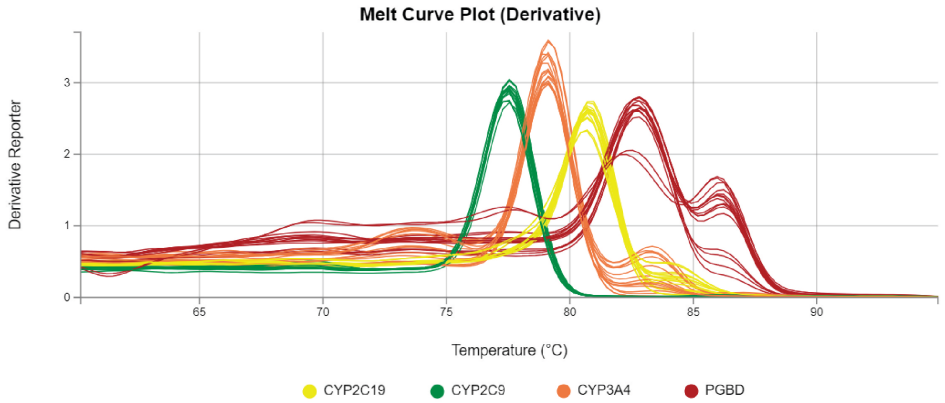


Figure 7. Melting curves of the four primers used for the qPCR using cDNA of HepaRG cells as template. *PGBD was used as housekeeping gene to normalize the Ct values measured for the other three targets: CYP2C9, CYP2C19, and CYP3A4.*

2.3 Single-cell RNA-seq experiments with primary human hepatocytes (PHHs)

Single-cell RNA-sequencing experiments with PHHs were performed following the illustrated experimental design below (**Figure 8**). The experiments explained here generated the data for the work that culminated in a co-first authored manuscript that can be found in Genome Biology [1]. Dr. Ioannis Deligiannis (Helmholtz Pioneed Campus, HMGU) participated in the experimental plan, experimentally co-executed, and designed the sequencing strategy of the first batch comprised of two donors with Lot Number HUM180812 and HUM4152.

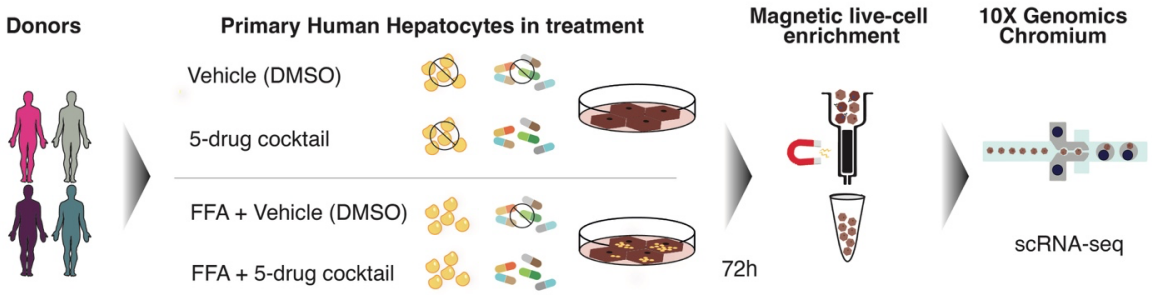


Figure 8. Schematic illustration of the experimental design followed to perform scRNA-seq experiments using commercially available cryopreserved PHH.

In brief, primary human hepatocytes from four donors were purchased in cryopreservation from Lonza. They were plated, preloaded with FFA for 6 h and incubated with either Vehicle (DMSO) or a 5-drug cocktail [285] for subsequent 66 h. After a total incubation time of 72 h, cells were collected, magnetic live cell enrichment was performed using a column and scRNA-seq was performed using the 10X Genomics Chromium.

Human liver donors

Commercially available Primary Human Hepatocytes (PHH) were purchased in cryopreserved vials from Lonza (Lonza, Walkersville, MD, USA) from four different donors (**Table 4**): HUM180812 (male, 57 years old, Hispanic) and HUM4152 (male, 18 years old, Caucasian), HUM181641 (male, 56 years old, Caucasian) and HUM4190 (male, 26 years old, Caucasian). All donors had a Body Mass Index in the normal range and were not diabetic.

Table 4. Summary of the four Primary human hepatocytes donor information obtained from Lonza.

Batch	Lot Number	Age	Ethnicity	Gender	BMI	Drug/Tobacco/Alcohol use	Diabetes	Thaw viability (%)	Thaw yield (Mio)
1	HUM180812	57	Hispanic	Male	21,1	No/No/Social	No	90,5	2,7
1	HUM4152	18	Caucasian	Male	24,3	Yes/No/No	No	82	6,9
2	HUM181641	56	Caucasian	Male	27	Yes/No/Social	No	88,6	88,6
2	HUM4190	26	Caucasian	Male	22	No/Yes/No	No	81	81

The PHHs used were cryopreserved, plateable, and interaction-qualified certified by Lonza with the catalog number #HUCPI. Lonza's internal quality control consists of the characterization of the hepatocytes regarding the three major drug-drug interactions, such as the enzymatic activity, transporter activity, and enzyme gene induction potential for CYP3A4, CYP2B6 and CYP1A2 (mRNA induction fold) [312]. These analyses are summarized in the Certificate of Analysis report provided with each vial. Therefore, these hepatocytes are considered metabolically competent concerning phase I and phase II metabolism, and suitable for their use in drug-related induction studies.

Each cryovial of PHH was thawed and plated according to Lonza's "Suspension and plateable cryopreserved hepatocytes: technical information and instructions.". The protocol was followed stepwise minutely, using the recommended thawing and plating media (Lonza, MCHT50 and MP250, respectively). The cells were dispensed and mixed using only wide orifice tips (Rainin, Ref. 17014297). For efficient cell seeding densities and attachment, cells were counted using Trypan Blue Exclusion Method and seeded into Collagen-I coated plates at a density of approximately 100,000 cells/cm^2 following the instructions provided by Lonza (Lonza, "Suspension and Plateable Cryopreserved Hepatocytes Technical Information and Instructions"). Six hours post-seeding, cells were washed with 1 mL of pre-warmed Maintenance Medium (Lonza, MCHT50) before the addition of treatment media. The treatment medium was renewed every 24h for a total incubation period of 72h post-seeding.

Free fatty acid incubation

Primary human hepatocytes were incubated with a mixture of the unsaturated oleic acid to the saturated palmitic acid in a 2:1 ratio at a concentration of 200 µM for 66 h, and a total of 72 h.

Free fatty acid medium preparation

Both palmitic and oleic acid were purchased in powder from Sigma Aldrich (P0500-10G and O1008-1G, respectively) and dissolved in ethanol to prepare a 50 mM stock solution [262]. In order to facilitate FFA uptake, pre-bounding of free fatty acids to 1% bovine serum albumin (BSA)

(10735078001, Sigma Aldrich) in a 1:5 molar ratio was performed by heating the mixture at 40ºC for 2 hours [262]. The FFA culture medium for cell incubation was prepared in a 50 mL Falcon tube containing 66,5 µL of palmitic acid, 133,5 µL of oleic acid, 500 µL of 10% BSA stock and 49,300 µL of pre-warmed PHH maintenance medium. Thereafter, the falcon tube was left open for 3 minutes under the laminar flow hood to evaporate the ethanol. Hence, a 200 µM mixture of a 2:1 ratio of the two free fatty acids was added to the cells in order to mimic the circulating levels in human steatosis [260].

Drug cocktail preparation and storage

The individual components of the 5-drug cocktail [285] were dissolved in sterile DMSO, filtered through a 0.2 µM syringe filter (Merck, SLGVV255F) and stored at -80ºC for a maximum of six months ("compound stock concentration"). The individual drugs were mixed at 200x concentration ("working concentration") and added to the cells to a final concentration of 80µM Caffeine (Sigma-Aldrich, Ref. 56396-100MG), 5µM Midazolam (LGC Chemicals, Ref. LGCFOR1106.00), 17µM Omeprazole (TRC Chemicals, Ref. 0635000), 20µM S-Warfarin (Sigma Aldrich, Ref. UC214-5MG) and 23µM Metoprolol (TRC Chemicals, Ref. M338815). The final DMSO concentration used on the cells was 0.5% (v/v) in all conditions.

Single-cell RNA-seq sample preparation

After a 72-hour incubation in treatment culture medium, cells were detached with prewarmed 0.25% Trypsin-EDTA (Gibco Life Technologies, 25200056) for 7 minutes, followed by the addition of an equal volume of pre-warmed maintenance medium to inactivate trypsin. The dissociated cells were then collected to pellet by centrifugation at 100 rcf for 5 min at room temperature (RT). Cells were washed twice with 1 mL of pre-warmed 1xPBS pH 7.4 (Gibco Life Technologies 10010023), followed by cells pelleting at 100 rcf for 5 min at RT. Live-cell enrichment was performed using the Dead-cell removal kit from MACS (Miltenyi Biotec) as follows: cells were pelleted at 100 g for 5 min at RT, resuspended in 100 µL of dead-cell removal microbeads and incubated for 15 min at RT. Dead cells were positively selected by passing the cell suspension through a MS column and performing a wash with a total of 2 ml of Binding buffer. Living cells were eluted from the column and collected in 2 mL Eppendorf tubes. After pelleting by centrifugation at 100 g for 5 min at RT, cells were resuspended in 1xPBS pH 7.4 supplemented with 0.04% BSA, stained with trypan blue to assess viability, and counted in a hemocytometer.

A single-cell suspension was obtained by dissociating cells with wide orifice pipette tips and preparing the target cell stock concentration for loading the 10X chip. Single-cell RNA-seq libraries were

prepared from each sample following the 10X Genomics Single Cell 3' Reagent Kit User Guide (v3 or v3.1, manual CG000183 and CG000204, respectively) in the Chromium Controller (10X Genomics). The quality control of cDNA and obtained final libraries was performed using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent). Library quantification was performed using the Collibri™ Library Quantification Kit (Thermo Fischer Scientific, A38524500) in a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher Scientific).

Single-cell RNA-seq library preparation and sequencing

Single-cell RNA-seq libraries were prepared from each sample following the 10X Genomics Single Cell 3' Reagent Kit User Guide (v3 for the first batch or v3.1 for the second, manual CG000183 and CG000204, respectively) in the Chromium Controller (10X Genomics). The first batch was prepared with Dr. Deligiannis. cDNA and obtained final libraries underwent quality control using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent). The quantification of the final libraries was performed using the Collibri™ Library Quantification Kit (Thermo Fischer Scientific, A38524500) in a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher Scientific) by Dr. Ioannis Deligiannis. Both batches were sequenced in a NovaSeq6000 sequencer (Illumina) (HMGU Core Facility for NGS Sequencing). The first batch (HUM180812 and HUM4152) performed by Dr. Ioannis Deligiannis was sequenced in an S2 flowcell at a depth of 250,000 reads per cell. The second batch (HUM181641 and HUM4190) performed by me was sequenced in an SP flowcell at a sequencing depth of 20,000 reads per cell. The sequencing length was set as indicated by 10X Genomics: 28/8/0/91.

Read alignment, counting, and filtering of the combined batches

The initial and exploratory analysis of the scRNA-seq data was performed by Xavier Pastor (Core Facility Genomics, Helmholtz Zentrum München) and myself. Thereafter, the follow-up computational analysis was performed by Ms. Maria Richter (Molecular ageing lab, Helmholtz Pioneer Campus, Helmholtz Zentrum München). When specified in the corresponding figure (Results, Chapter III), the results were computed by Ms. Maria Richter, otherwise, analysis and figures were performed by myself.

Reads were aligned to GRCh38 and counted using 10X Genomics Cellranger 4.0.0 with standard parameters set individually for each batch due to their different sequencing depths. The individual count matrices were combined into a common count matrix consisting of 63,527 cells and 19,971 genes. Cells with at least 1,000 counts and 500 genes were kept. Genes present in at least 5 cells and had fewer than 5 million reads were kept. The package *Scrublet* [313] was used to identify doublets. Due to different ploidy statuses among hepatocytes, a lenient cutoff of 0.15 was set to avoid unintended

removal of tetraploid cells, leading to 1.7 % doublets removed. Cells with mitochondrial reads over 1 %were removed, which resulted in a filtered matrix of 49,378 cells times 16,256 genes.

Transcriptional variability calculation through the coefficient of variation

Lowly expressed genes have generally higher transcriptional variability, and they were filtered out. Genes with a mean normalized log-transformed expression greater than 0.25 were kept and the coefficient of variation per condition in each of the subgroups was calculated on 3,434 genes. The coefficient of variation on normalized and log-transformed counts was calculated using the formula described by Canchola, 2017 [314], where σ^2 is the variation of gene *j* in the group of interest.

$$CV = \sqrt{e^{\sigma^2}} - 1,$$

In every subgroup, a MannWhitneyU test was performed to validate if the coefficient of variation differed significantly between treatment conditions or ages.

Single-nucleotide polymorphisms (SNPs) identification

The data for the SNPs detection was obtained from the FASTQ files of the four sequenced scRNA-seq (10X Genomics) samples from each of the two donors. These samples corresponded to cells treated with DMSO, Cocktail, FFA+DMSO, or FFA+Cocktail per donor.

Long Ranger (10X Genomics) was used to call SNPs from each of the four samples comprised per donor. For that, the function *"longranger wgs"* was used, aligning reads to the GRCh38-2.1.0 reference human genome. *Longranger wgs* uses the Genome Analysis Toolkit (GATK) [315] tool to phase and call SNPs. In this case, *"gatk"* version 4.0.3.0 was used. An illustrative example of the settings used running "*longranger* wgs" on one of the eight samples (Donor 1, SampleMUC14324) is shown here:

```
/home/esanchez/gatk-4.0.3.0/longranger-2.2.2/longranger wgs --id=SampleMUC14324
--sex=male --
fastqs=/home/hpc/martinez/00_projects/human/precision_toxicology/rep1/fastq/2001
14_A00623_0079_BHMK2NDMXX_lanes1_2/Project_all/ --indices=SI-GA-E1 --
reference=/home/esanchez/gatk-4.0.3.0/refdata-GRCh38-2.1.0 --sample=MUC14324 --
vcmode=gatk:/home/esanchez/gatk-4.0.3.0/gatk-package-4.0.3.0-local.jar --
uiport=3600
```

Only SNPs detected in all four samples of one donor were kept to ensure they were consistently detected in all cells analyzed from that specific donor.

The resulting ".vcf" files were converted into a dataframe using the *scanpy* function *"allel.vcf_to_dataframe"*. The SNPs called in the chromosomes where the gene of interest is located were displayed using the function "*df_phased.loc*" and saved in a ".csv" file per each of the four samples. Then, the four .csv files were combined into a single one, and the SNPs detected four times samples were kept by using the command line function *gawk.* The resulting SNPs were contrasted in the PharmVar [316], ClinVar [317, 318] and NCBI databases (dbSNP) [319] to examine their incidence in the population and clinical relevance. **Table 5** summarizes the donor, ethnicity, the cytochrome of interest, the chromosome where the cytochrome gene is located, the position, the variant detected, the potential alleles, the SNP ID named rsID, the type of variant, and the NCBI link to the corresponding SNP information.

Table 5. Summary of detected SNPs in the two human donors. *Link to the NCBI database dbSNP for the orresponding SNP can be found in the last column.*

Donor	Ethnicity	CYP	Chromosome	Position	Variant	Alleles	rsID	Type	Information
1	Caucasian	CYP2C9	**chr10**	94942290	T	C>T	CYP2C9*2 - rs1799853	Missense variant	https://www.ncbi.nlm.nih.gov/clinvar/RCV000008920.2/
1	Caucasian	CYP2C9	**chr10**	94942538	A	G>A	Rs2860905	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs2860905
1	Caucasian	CYP2C9	**chr10**	94947445	T	C>T	rs4086116	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs4086116#publications
1	Caucasian	CYP2C9	**chr10**	94950236	C	T>A or T>C or T>G	rs2984310	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs2984310
1	Caucasian	CYP2C9	**chr10**	94952643	G	A>G	rs2475376	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs2475376#publications
1	Caucasian	CYP2C9	**chr10**	94972974	G	T>A or T>G	rs1856908	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs1856908
1	Caucasian	CYP2C9	**chr10**	94982060	G	A>C or A>G	rs1934968	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs1934968#publications
1	Caucasian	CYP2C19	**chr10**	94804000	A	G>A	rs4494250	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs4494250
1	Caucasian	CYP2C19	**chr10**	94821337	G	A>C or A>G	rs10786172	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs10786172#publications
1	Caucasian	CYP2C19	**chr10**	94842866	G	A>C or A>G	Rs3758581	Missense variant	https://www.ncbi.nlm.nih.gov/snp/rs3758581
2	Hispanic	CYP1A2	**chr15**	74755085	C	T>C	rs2470890	Intronic mutation	https://www.ncbi.nlm.nih.gov/snp/rs2470890#publications
2	Hispanic	CYP2C9	**chr10**	94952643	G	A>G	Rs2475376	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs2475376#publications
2	Hispanic	CYP2C9	**chr10**	94981151	C	T>C	rs9332197	Intronic variant	https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query=rs9332197

2	Hispanic	CYP2C9	**chr10**	94982060	G	A>G or A>G	rs1934968	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs1934968#publications
2	Hispanic	CYP2C19	**chr10**	94821337	G	A>C or A>G	rs10786172	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs10786172#publications
2	Hispanic	CYP2C19	**chr10**	94842866	G	A>C or A>G	Rs3758581	Missense variant	https://www.ncbi.nlm.nih.gov/snp/rs3758581
2	Hispanic	CYP2C19	**chr10**	94849811	C	T>C	rs4917623	Intronic variant	https://www.ncbi.nlm.nih.gov/snp/rs4917623
2	Hispanic	CYP2D6	**chr22**	42126722	T	G>T	rs79392742	Missense variant	https://www.ncbi.nlm.nih.gov/snp/rs79392742

An example of the NCBI report generated regarding the SNP of interest with rsID rs1799853 is shown in **Figure 9**.

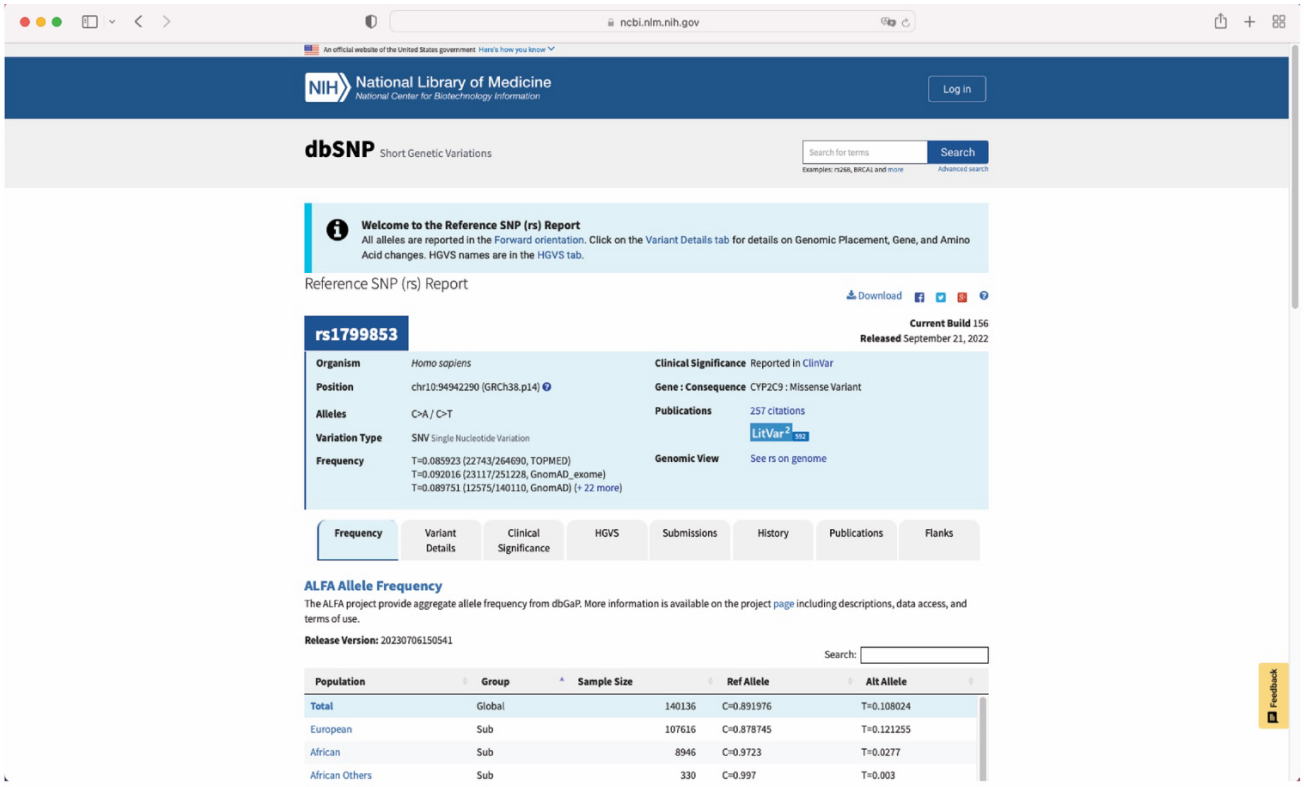


Figure 9. NCBI report of the detected SNP with rsID rs1799853. *Browser screen capture of the results after searching for rs1799853 in the dbSNP database from the NCBI [319].*

After finding the variant in the dbSNP database [319], the clinical significance of it was displayed and reported in ClinVar [317, 318] together with a link to the ClinVar website. By clicking on it, ClinVar the variation, the gene affected, the protein change, the condition, and the clinical significance are shown (**Figure 10**).

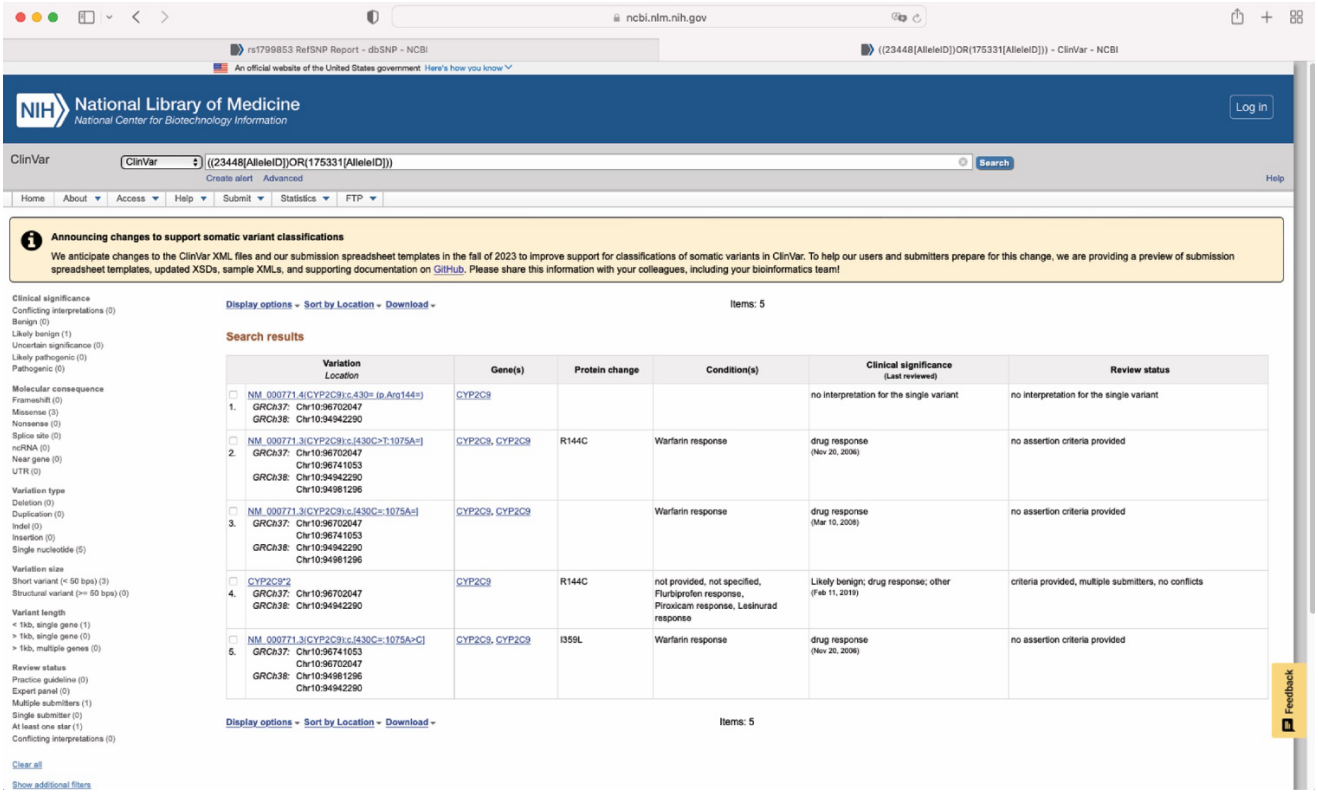


Figure 10. ClinVar database screen capture after clicking on the dbSNP link. *The chromosome, position, protein change, condition, clinical significance, and gene regarding the SNP of interest (rsID 1799853) can be seen.*

According to the SNPs detected using Long Ranger, the variant fount was a thymidine (T) replacing a cytosine (C) nucleotide in position 94942290 of chromosome 10, corresponding to the CYP2C9 gene. Since the alignment was performed using the GRCh38 reference genome, the image below shows the exact variant that was detected (**Figure 11**).

Figure 11. Variant detected using Long Ranger found in ClinVar, *Screen capture of the corresponding detected variant in the CYP2C9 gene, substituting a thymidine by a cytosine nucleotide.*

When clicking on the link, further information and details about the variant of interest were displayed on the ClinVar website, including literature reports associated to the information on this SNP (**Figure 12**).

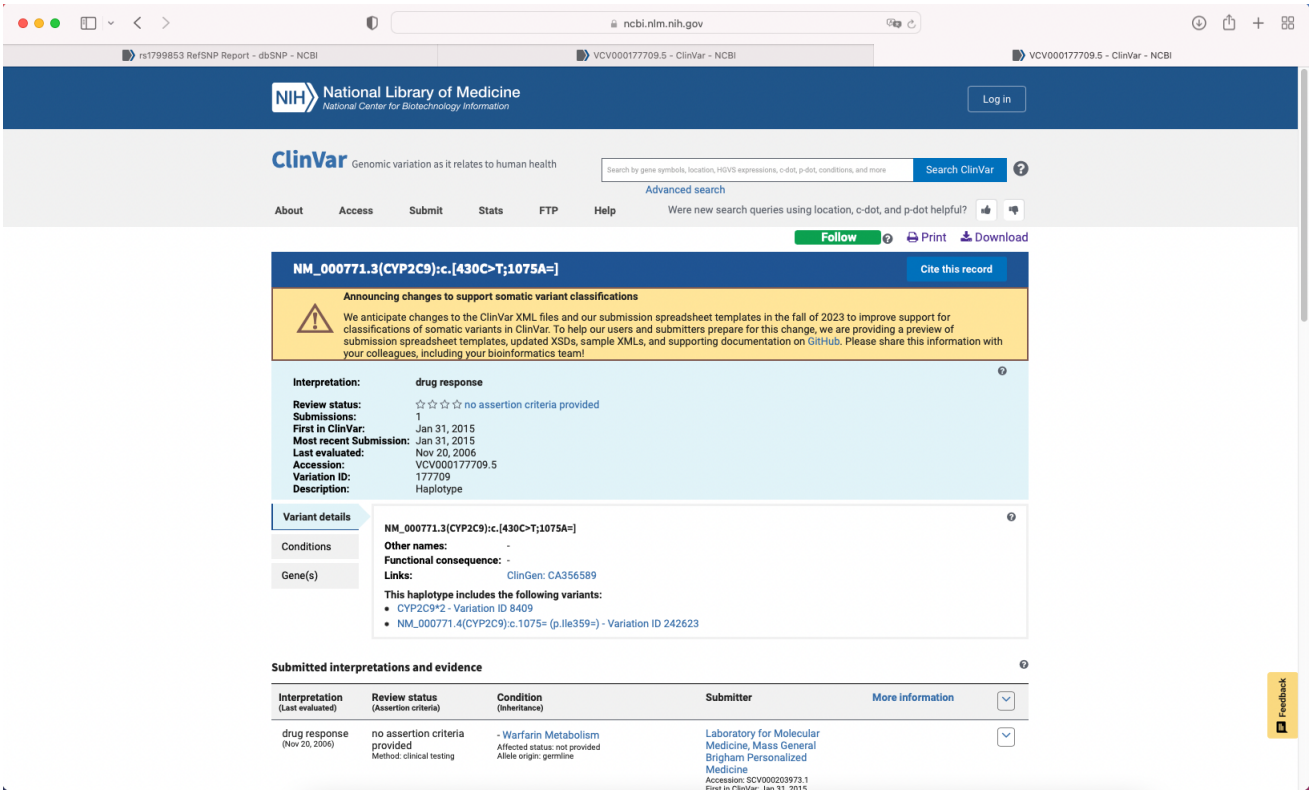


Figure 12. Further details and information about the SNP of interest are displayed in ClinVar. *Screen capture showing the variant details including literature reports.*

The same procedure was performed to find SNPs present in the four samples from the same donor for each of the five targeted cytochromes (*CYP1A2, CYP2C9, CYP2C9, CYP2D6* and *CYP3A4*).

Marker gene list

A curated list of 660 known human marker genes was curated by me performing thorough literature research on the genes composing the list and published as "Supplementary Table 2" in the manuscript Sanchez-Quant, Richter *et al.* 2023 [1]. The genes belonged to different biological categories: amino acid metabolism, aging, anti- and pro-apoptosis, autophagy, bile synthesis, carbohydrate metabolism, circadian rhythm, downregulated in PHHs after 48 h in culture, downregulated in PHHs after 72 h in culture, ER stress, lipid metabolism, NAFLD-related, G1 phase, G1-S phase, G2-M phase, S phase, hepatocellular carcinoma, hepatoblasts, hormone metabolism, inflammation, lipid droplet formation, maintained in expression in PHHs after 48 h in culture, mature hepatocyte, liver-enriched nuclear receptors, organoid differentiation, oxidative stress, phase I metabolism of eicosanoids, phase I metabolism of fatty acids, phase I metabolism of sterols, phase I metabolism of vitamins, phase I metabolism of xenobiotics, phase II metabolism enzymes, phase III metabolism transporters, retinoid metabolism, steroid hormone metabolism, stress, and toxicity.

Subgroups of hepatocytes were identified using this curated marker gene list (performed by Ms. Maria Richter). Thereafter, *scanpy* and *episcanpy* functions were used for further downstream analysis and figure generation.

2.4 Experimental optimization of the scATAC-seq approach for the study of PHHs

An aim of this thesis is to study upstream molecular events affecting gene expression in individual primary human hepatocytes by assessing chromatin accessibility across the genome. Hence, the Assay for Transposase-Accessible Chromatin with high-throughput sequencing at single-cell resolution (scATAC-seq) was applied. The stepwise protocol is depicted in the schematic representation in **Figure 13**, where the different steps were optimized for the nature of hepatic cells, which are voluminous cells and possess a stiff cytoplasmic membrane.

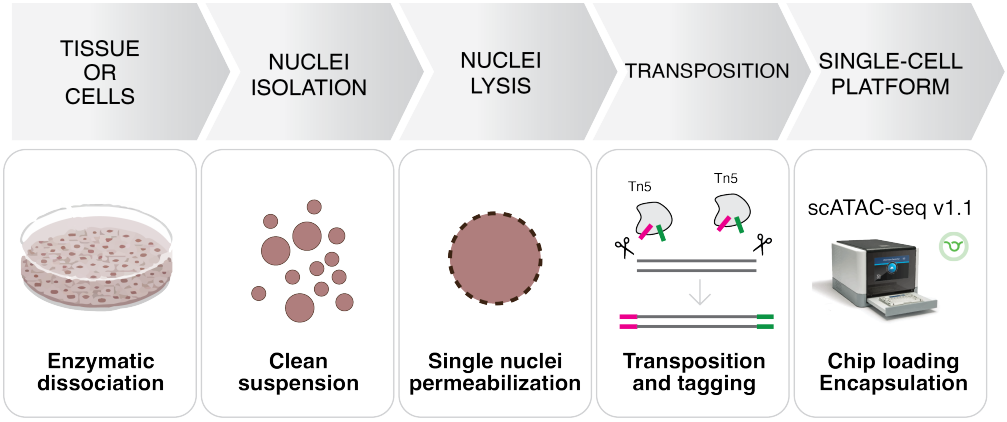


Figure 13. Schematic representation of the stepwise procedure for scATAC-seq.

Two great challenges presented in this pipeline are therefore a) the obtention of a clean suspension of singly dissociated nuclei, and b) the nuclear membrane permeabilization to open pores without completely disintegrating its integrity. Here, I will explain the optimization of specifically those two steps in fresh and frozen murine liver tissue and in primary human hepatocytes in bulk analyses. The optimized conditions were thereafter used on PHHs to perform scATAC-seq using the 10X Genomics Chromium platform.

a) Bulk experiments. Optimization of nuclei isolation for further scATAC-seq on PHHs

i) Flash-frozen and fresh mouse liver tissue – Swelling buffer nuclei isolation and nuclei lysis

The here used hypotonic swelling buffer was provided by the authors in the Chen *et al*. 2020 publication, constituting the first scATAC-seq publication in fresh murine liver tissue [48]. For the optimization purposes, the buffer was tested in both snap-frozen and fresh tissue from wild-type C57BL/6 mice.

A piece of flash-frozen liver tissue was chopped in a cold petri dish on ice in 2 mL of prechilled Swelling buffer (10 mM Tris-HCl pH7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$) and transferred into a pre-chilled douncer homogenizer. Subsequently, the suspension was dounced 10 times with the loose pestle (A), incubated 20 min on ice and dounced 20 times with the tight pestle (B). The suspension was filtered through a 40 µM CellTrics® cell strainer and centrifuged at 300 g for 5 min at 4 ºC in a swing bucket centrifuge. The pellet was resuspended in 1 mL of prechilled swelling buffer with 10% glycerol. Slowly and drop-wise, 500 µL of cold Swelling Lysis Buffer (swelling buffer + 10 % glycerol + 1 % NP-40) were added with occasional flicking. The suspension was incubated on ice for 5 min and then span down at 500 rcf for 5 min at 4ºC. The supernatant was discarded and the pellet obtained was slowly resuspended in 1 mL of Wash buffer from 10X Genomics nuclei isolation protocol composed of 10 mM Tris-HCl pH 7.4, 10 mM NaCl, $MgCl_2$, 1% BSA, 0.1 % Tween-20 and nuclease-free water [68]. The cells were counted and 50,000-100,000 cells taken out for further processing. The nuclei were then incubated in Lysis Buffer from 10X Genomics [68] containing 10 mM Tris-HCl pH 7.4, 10 mM NaCl, $MgCl_2$, 1% BSA, 0.1 %Tween-20, 0.01 % Digitonin, 0.1% NP-40 and nuclease-free water for 10 min on ice. After 10 min, 1 mL of cold Wash Buffer was added and slowly pipette-mixed 5 times, centrifuged down at 500 g for 5 min at 4 ºC in a swing bucket rotor. Then, the supernatant was carefully removed leaving 100 µL to be removed with a p20.

Tagmentation (bulk experiments) and clean-up

For the experiments performed as a bulk, the tagmentation mix was prepared as described by Corces *et al.* 2017 [320]. A total of 50 µL for one reaction were prepared composed of 2,5 µL of Tn5 transposase (Illumina, Nextera kit, Illumina Cat No. FC-121-1030), 25 µL of 2X TD Buffer (Illumina,), 16.5 µL of PBS, 0,5 µL of Tween-20 (Sigma Aldrich 11332465001) and 0,5 µL of Digitonin (Promega, G9441). The reaction tube was placed in a thermomixer for 30 min at 1000 rpm mixing.

Once finished, the tagmented DNA was cleaned up using a ZYMO DNA clean-up kit adding 250 µL of DNA Binding Buffer, a fixed angle rotor and 10,500 rcf for the centrifugation steps. The membrane was incubated with 20 µL Elution Buffer for 5 min and the DNA eluted at 13,500 rpm. The amplification PCR Master mix was prepared using Nextera Primer 1 (I5 index), Nextera primer 2 (I7 index) and 2X NEB Next Master mix for one reaction. A total of 10 µL of tagmented and cleaned DNA were used together with 15 µL of master mix. The PCR was set for: 5 min at 72 °C, 30 sec at 98

ºC, [10 s at 98 °C, 30 s at 63 °C, 1 min at 72 °C] × 18 or 16X, 4 ºC infinite hold. The obtained libraries were assessed using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent) in a Bioanalyzer 2100.

ii) Cryopreserved primary human hepatocytes nuclei isolation – homogenization buffer, swelling buffer and mechanical dissociation

Cryovials were thawed in the water bath at 37 ºC leaving the lid afloat for 2 min, then disinfected and poured into a 50 mL Falcon tube containing Thawing Medium (Lonza, MCHT50). To rinse the cryovial, 1 mL was used. Then, the falcon was rocked slowly and carefully and the cells pelleted at 100 g for 8 min at room temperature. The supernatant was then discarded and the cells resuspended in 1 mL of Maintenance Medium + supplement (Lonza, MM250-1 and -2). Then, 2 million cells were transferred to a clean Eppendorf tube and washed twice with 500 cold PBS.

Cells were centrifuged down at 500 rcf for 5 min at 4 ºC and resuspended in 1 mL of homogenization buffer containing 250 mM Sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris buffer pH 8.0, 1% DTT, one tablet protease-inhibitor (cOmplete EDTA-free, Life Technologies, 1187358001), 0.3 % Triton-X and 0.2 % NP-40. Using wide-bore pipette tips, cells were transferred to a 2 mL douncer homogenizer, 5 strokes performed with the loose pestle (A), incubated on ice for 10 min and 25 strokes performed with the tight pestle (B). The nuclei were inspected under the microscope to validate that the cell membrane was disrupted and transferred to a 2 mL Eppendorf tube. The douncer was washed with 400 µL of HB and all together centrifuged at 500 rcf for 5 min at 4 ºC. Thereafter, the supernatant was removed and the pellet was resuspended slowly in 1 mL of Swelling Buffer (10 mM Tris-HCl pH7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$) and 1 mL of prechilled Swelling Buffer with 10% glycerol was added dropwise and then pipette-mixed 5 times. The mixture was incubated for 10 min on ice with occasional flicking. Then, 1 mL of 10X Genomics Wash Buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, $MgCl_2$, 1% BSA, 0.1 % Tween-20 and nuclease-free water), incubated on ice for 10 min and 60,000 taken out for further steps.

To lyse the nuclei, these were centrifuged at 500 rcf at 4 ºC for 5 min. Then, the supernatant was discarded, and nuclei resuspended in 200 µL of 10X Genomics Lysis Buffer with 0.2 % NP-40 instead of the original 0.1 % concentration (10 mM Tris-HCl pH 7.4, 10 mM NaCl, $MgCl_2$, 1% BSA, 0.1 %Tween-20, 0.01 % Digitonin, as a minor 0.2% NP-40 and nuclease-free water). Nuclei were lysed for 10 min on ice and then 1 mL of Wash Buffer was added followed by rocking the tube gently. Nuclei were pelleted at 500 rcf for 5 min at 4 ºC, then 500 µL of PBS added without resuspending and nuclei spun down at 500 rcf for 5 min at 4 ºC.

Tagmentation and clean-up

For the experiments performed on PHHs, the tagmentation mix was prepared as per in the 10X Genomics demonstrated protocol [68]. Nuclei were resuspended in 500 µL of 1X Nuclei Storage Buffer, then centrifuged down at 500 rcf for 5 min at 4 ºC and the supernatant was discarded. Then, the transposition mix was prepared with 7 µL of ATAC Buffer and 3 µL of ATAC enzyme [321]. Then, using the Nuclei Concentration Guidelines, a total of 5 µL of the mixture of nuclei (2.9 µL) and diluted Nuclei Storage Buffer (2.1 µL) were incubated in a thermocycler for 60 min at 37 ºC with the lid temperature at 50 ºC. Thereafter, the libraries were assessed using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent) in a Bioanalyzer 2100.

b) Single-cell ATAC-seq experiment on PHHs

Nuclei isolation

For the experiment on PHHs at single-cell resolution, nuclei were isolated as explained above in section 2.2.2, part A), paragraph ii).

Human liver donors for single-cell ATAC-seq experiment

Cryopreserved commercially available Primary Human Hepatocytes (PHHs) were purchased from Lonza from two different donors: HUM180812 (male, 57 years old, Hispanic) and HUM4152 (male, 18 years old, Caucasian). These donors were as well used for scRNA-seq experiments. They had a BMI classified in the normal range and were not diabetic (**Table 4**, first two rows) Lifestyle-related habits such as tobacco, alcohol or drugs consumption were not present or was defined as social in the case of alcohol.

Single-cell ATAC-seq library preparation and sequencing

After GEM incubation the single-cell ATAC-seq libraries were prepared from each sample following the 10X Genomics Chromium Next GEM Reagents Kit Reagent Kit User Guide (v1, manual CG000209, RevD) in the Chromium Controller (10X Genomics). cDNA and obtained final libraries underwent quality control using a Bioanalyzer High Sensitivity DNA Analysis assay (Agilent). The quantification of the final libraries was performed using the Collibri™ Library Quantification Kit (Thermo Fischer Scientific, A38524500) in a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher Scientific).

The libraries were sequenced in an SP 100 flowcell in a NovaSeq 6000 sequencer (Illumina) (HMGU Core Facility for NGS Sequencing). The sequencing length was as indicated by 10X Genomics (50, 8, 16, 50) for scATAC sequencing. These conditions yielded a depth of approximately 50,000 reads per nucleus.

Read alignment, peak calling and count matrix generation, filtering and quality control

Reads were aligned to GRCh38 by Mr. Patrick Hanel (Institute of Computational Biology, Helmholtz Zentrum München) and counted using 10X Genomics Cellranger ATAC 4.0.0. Signac was used for the QC of the samples, to plot the TSS enrichment and the nucleosome periodicity. Then, *MACS2* [322] was used to identify peaks within 5 kb upstream of the transcription start site (TSS) in each of the eight samples individually. Peak-based count matrices constructed per sample were converted into *Seurat* objects using *SeuratDisk* and a chromatin assay object created using the fragments file from Cell Ranger ATAC. G ranges were obtained from EnsDb.Hsapiens.v86 and subsequently, TSS enrichment and nucleosome signal calculated and plotted using Signac [376].

To merge samples into the same feature space for further downstream analysis, the generated peaks were combined using *Bedops* [323] in a single peaks file. The merged peaks were used to build a count matrix per sample using the *Episcanpy* function "epi.ct.bld_mtx_bed" [324]. Subsequently, the eight count matrices were concatenated for further joint processing. For filtering, cells with a minimum of 100 features and a minimum of 50,000 highly variable features were kept. Filtering resulted in a joined matrix consisting of 59,142 observations times 50,244 variables used for downstream analysis.

Regression, clustering and doublet detection

Due to the different coverage between samples, the number of features was a confounding variable and they were therefore regressed out using the *Scanpy* function "sc.pp.regress_out". Using *Episcanpy* functions, PCA was calculated using the top 100 components and 15 neighbors, and embedding was performed using PCAs. *Scrublet* [313] was used to identify doublets, setting the expected doublet rate to 6.2%, as indicated by 10X Genomics when the targeted recovery is 5,000 cells [321]. *Episcanpy* functions (0.2.0) were applied for downstream analysis.

3. Results

I. Block I (Chapters I and II): Single-cell omics for dissecting cellular heterogeneity in complex mouse tissues using plate-based approaches.

Plate-based single-cell omics approaches are a low-throughput method offering an advantage in specific scenarios, for instance, when studying the transcriptome of rare cell subpopulations such as oxytocin-positive hypothalamic neurons obtained from flash-frozen mouse samples. In this first block of my dissertation, a plate-based single-cell transcriptomic approach such as snRNA-seq2 [19] methodology was applied to explore the transcriptomic responses of murine hypothalamic oxytocin-positive neurons. Specifically, their susceptibility to showing altered transcriptomic responses upon an obesogenic high-fat high-sugar diet.

Plate-based approaches are not only suitable for transcriptome analysis but also for studies at the gene regulatory level. The sensitivity offered by a plate-based methodology can be of advantage to explore the differential chromatin accessibility across different conditions. Coupled with FACS sorting nuclei based on their genome content, the ploidy levels of hepatocytes can be studied. In this thesis, the initial steps for a plate-based approach to enable the analysis of the open chromatin landscape of murine diploid (2n) vs. tetraploid (4n) hepatocytes were developed.

CHAPTER I: Investigating the heterogeneity among oxytocin-expressing neurons. The snRNA-seq2 methodology in brain vs. liver tissue.

Small brain regions such as the hypothalamus yielding a low number of cells, hinder the application of high-throughput methodologies, especially when interested in deeply characterizing rare subpopulations [325]. These require millions of cells as input, making low-throughput methods combined with FACS purification an ideal choice to enable the recovery and analysis of all available cells. In addition, the study of cells of the Central Nervous System (CNS) at single-cell resolution is in general challenging, in part due to the difficulty of isolating intact whole cells [64]. Neurons are highly intertwined, making considerable damage necessary to dissociate them by physical means, such as laser capture microdissection [64].

The recently developed single-nucleus RNA-seq2 methodology, (snRNA-seq2) based on the SMART-seq2 chemistry provides a robust pipeline for the analysis of full-length transcripts with an improved nuclei lysis achieved in two consecutive steps [19] (Methods). This medium-throughput method was developed using mouse flash-frozen liver as the model organ of study; however, its application potentially reaches a wide range of tissues, cells and experimental set-ups. To interrogate the role of oxytocin-producing hypothalamic neurons in feeding homeostasis and behavior, mice were fed *ad libitum* with chow standard (CS) diet or an obesogenic high-fat, high-sugar (HFHS) diet for more than 12 weeks, provided by Dr. Tim Gruber (Institute for Diabetes and Obesity, Helmholtz Zentrum München). Then, nuclei from the mouse hypothalamus were isolated by Dr. Raian Contreras (Institute for Diabetes and Obesity, Helmholtz Zentrum München) and nuclei of GFP^+ neurons were FACS-purified [37, 306] to perform transcriptomic profiling using the newly emerged snRNA-seq2 improved lysis buffer methodology (Methods). After sorting GFP^+ nuclei, the library preparation and sequencing were performed and quality control of the libraries electropherograms performed using a Bioanalyzer 2100 system (**Figure 14**).

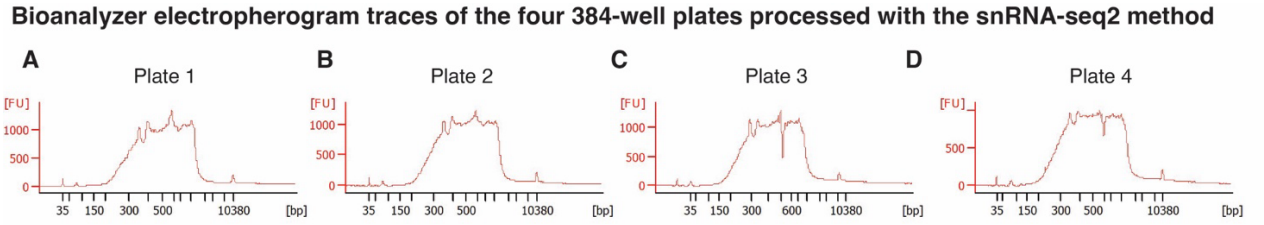


Figure 14. Bioanalyzer electropherograms of the four libraries prepared. ***(A)*** *Profile of plate 1.* ***(B)*** *Profile of plate 2.* ***(C)*** *Profile of plate 3.* ***(D)*** *Profile of plate 4.*

An initial count matrix of 1,536 single nuclei and 55,579 genes was generated by Dr. Viktorian Miok (Institute for Diabetes and Obesity, Helmholtz Zentrum München). Filtering and normalization resulted in a count matrix of 1,202 nuclei times 13,867 genes (Methods). A median number of 2,304 genes were detected, with approximately 80,000 counts assigned to ERCCs (**Figure 15**).

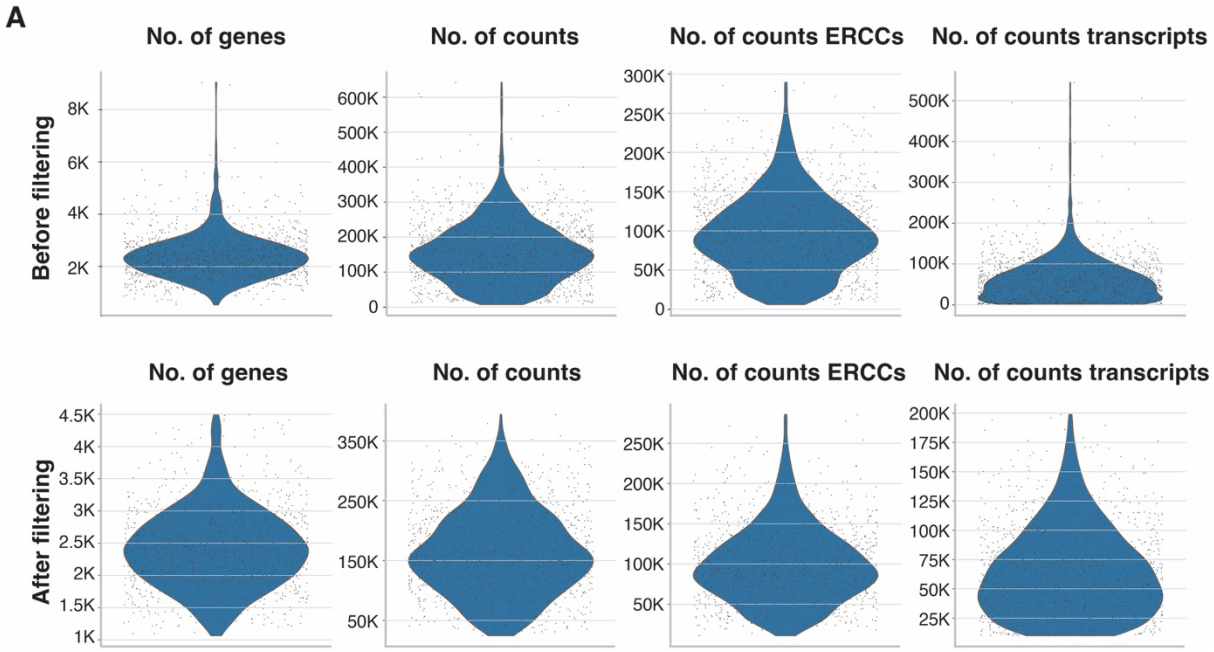


Figure 15. Quality control was performed on the nuclei isolated from OXT+ neurons. ***(A)*** *Number of genes, number of counts, number of counts assigned to ERCCs and number of counts assigned to transcripts prior to filtering and normalization;* ***(B)*** *Number of genes, number of counts, number of counts assigned to ERCCs and number of counts assigned to transcripts after to filtering and normalization. Graphs and QC computed by Dr. Viktorian Miok, HMGU.*

The snRNA-seq2 methodology was initially developed using mouse liver tissue. Here, this methodology was applied to nuclei from hypothalamic OXT^{+} neurons. The number of genes detected in both liver and brain were compared, to assess its sensitivity in different complex tissues and cell types. It is necessary nonetheless to consider that the liver physiology and physical characteristics differ largely from those of other tissues such as brain for several reasons, one of them being the larger size of the hepatocytes (parenchymal cells), ranging around 40 µM diameter [326, 327]. In liver tissue, an average of 3,400 genes per nucleus was detected, whereas from flash-frozen mouse hypothalamus, the mean number of detected genes was lower, with 2,304 genes per nucleus on average (**Figure 16**, Methods). An explanation for this could be that the cell size increases with the ploidy status and it has been shown in mouse and human liver that the nuclei volume of hepatocytes approximately doubles with doubling of the DNA content [95, 328, 329]. Transcript abundance has been shown to correlate with cell volume at single-cell level due to an increase in overall transcription in larger cells [330, 331]. Therefore, a lower transcriptome could be expected from brain nuclei vs. liver nuclei [10].

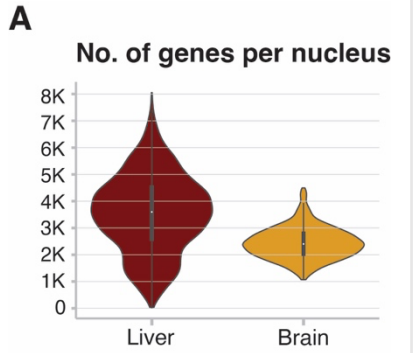


Figure 16. Number of genes per nucleus detected on average. ***(A)*** *Comparison of liver tissue, where the snRNA-seq2 method was developed vs. brain hypothalamus.*

Moreover, in a study performed using a modified version of Smart-seq2 by Batiuk et al., an average of 2,100 genes per cell were detected [332]. In a study performed by Thrupp *et al.* [333], 1,791 genes

per nuclei and 2,242 genes per cell were detected from human brain samples, which stay below the 2,304 genes per nucleus captured in this study (**Figure 16**). Chen *et al.* reported more than 2,000 genes per cell detected in their murine hypothalamus scRNA-seq characterization study [79]. This highlights the superior sensitivity and capability of this approach to characterize the nuclear transcriptome of scarce neuronal subpopulations.

Next, to visualize how similar the transcriptomic profile was among the nuclei in the study, Principal Component Analysis (PCA) was calculated after using *scran* [334] for normalization and *ComBat* [335, 336] for batch effect correction. The number of genes and counts detected were homogeneous throughout all cells. A separation between nuclei isolated from animals fed on a chow standard (CS) diet vs. HFHS-fed mice was not observable (**Figure 16**).

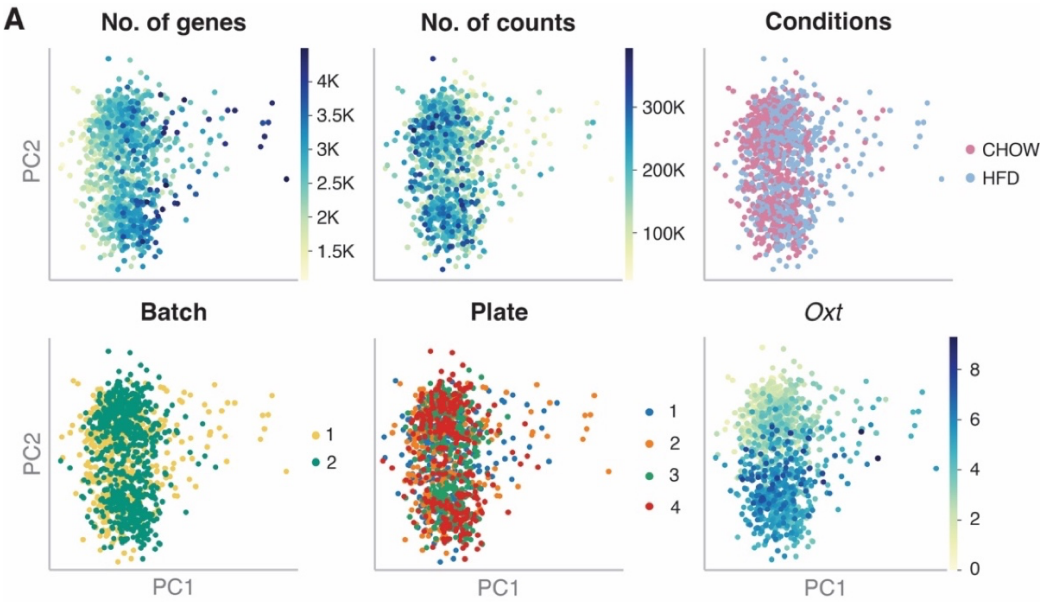


Figure 17. Homogeneous number of genes and number of counts throughout experiments and treatment diets. ***(A)*** *PCA showing the number of genes, number of counts, dietary conditions, batches and plates.* ***(B)*** *PCA plot depicting oxytocin expression across cells after normalization using scran and batch correction using combat.*

Purification of GFP-oxytocin-expressing neurons was performed by FACS-sorting nuclei into 384-well plates, therefore *Oxt* expression was expected to show a high level across all nuclei in the study. This was confirmed by observing the expression of the oxytocin (*Oxt*) gene across all cells in the PCA representation (**Figure 17**).

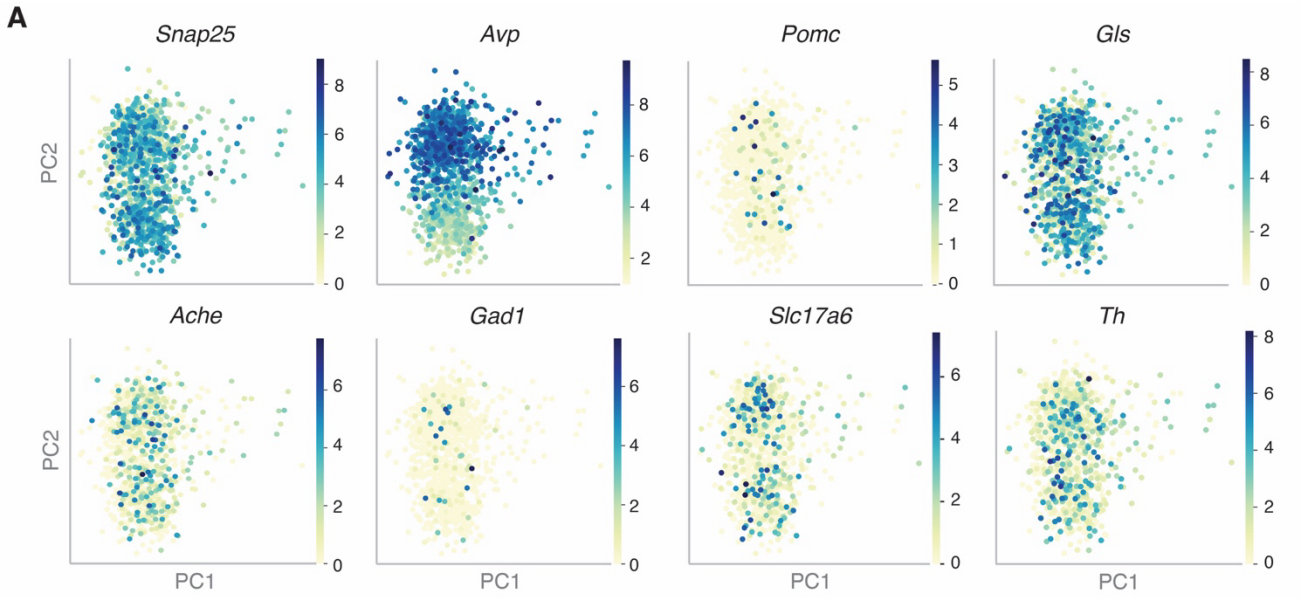


Figure 18. Oxytocin-positive neurons show coexpression of marker genes for different neuronal types. *(A) PCA plot depicting the expression of marker genes Snap25, Avp, Pomc, Gls, Ache, Gad1, Grin2 and Th across hypothalamic OXT⁺ neurons.*

Additionally, markers of different types of neurons were distinguished such as pan-neuronal markers like *Snap25* [337]. Arginine vasopressin *Avp,* which is known to antagonize oxytocin-mediated functions [338] was detected in the nuclei that lowly expressed *Oxt* and vice versa, indicating that the expression of these two neuropeptides is complementary and specific to two different neuron subtypes. The prohormone proopiomelanocortin (*Pomc*), known to play a key role in feeding behavior and energy homeostasis [339] could only be identified in a small number of nuclei, indicating the expression of both *Oxt* and *Pomc* in those (**Figure 18**). A set of nuclei was expressing tyrosine hydroxylase (*Th*), and it has been shown that these neurons are glucose homeostasis modulators in mice [340]. Furthermore, markers for different types of excitatory neurons were detected, among which are glutamatergic neuron markers such as *Slc17a6* [341], glutaminergic neurons such as *Gls* [342], cholinergic neurons *Ache* [343]. The inhibitory GABAergic neurons marker *Gad1* [344], was detected solely in a minor fraction of cells (**Figure 18**), indicating a predominant excitatory activity in the nuclei of the neurons in the captured cells. This indicates that the *Oxt* is co-expressed with several neuronal subtype marker genes in the OXT^+ sorted nuclei.

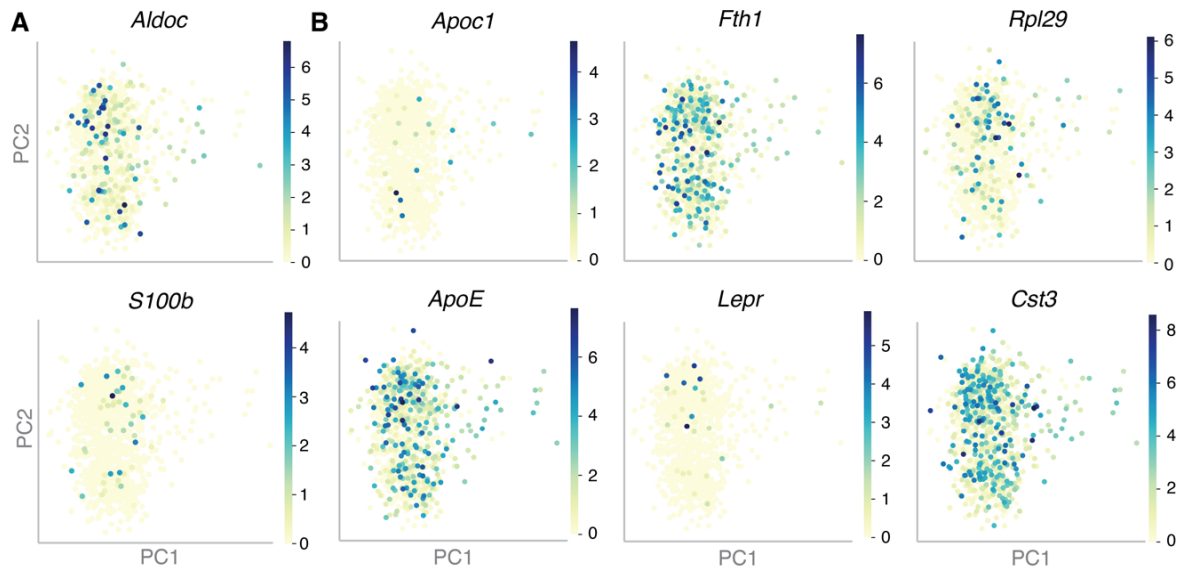


Figure 19. Gene expression of astrocyte marker genes and microglial activation-related genes. ***(A)*** *PCA showing astrocyte marker genes Aldoc and S100b.* ***(B)*** *PCA depicting microglial activation-related genes Apoc1, Fth1, Rpl29, ApoE, Lepr and Cst3.*

Marker genes, such as *Aldoc* or *S100b* have been used as astrocyte markers [345, 346], and they were detected in a minor fraction of oxytocin-expressing neurons in this dataset (**Figure 19**). This suggests that there is heterogeneity in the gene expression patterns found within the OXT^{+} population of neurons, indicating that they may respond differently to external stimuli, such as exposure to a high-fat/high-sugar diet.

Furthermore, the snRNA-seq2 method allowed for the detection of microglial activation-related genes such as *ApoE*, *Fth1*, *Cst3*, and *Rpl29, Apoc1* and *Lepr* (**Figure 19**), which were found to be depleted in the nuclear vs. whole cell transcriptomics dataset generated using the droplet-based snRNA-seq from 10X Genomics Chromium published by Thrupp *et al.* [333]. This showcases the high sensitivity, adaptability, and reproducibility of the snRNA-seq2 method to be used in other tissues and experimental setups.

Taken together, these results show that the snRNA-seq2 method allows for the exploration of complex tissues such as brain, as well as for the interrogation of the gene expression profiles of single hypothalamic nuclei from flash-frozen samples. Subtypes of oxytocin-expressing neurons were found, showing mutually exclusive marker gene expression that may lead to distinct responses toward stimuli such as an obesogenic diet.

CHAPTER II: Exploring the gene regulatory landscape of rare populations at single-cell level using a low to medium throughput approach.

In this chapter, the initial steps for the development of a low-throughput plate-based scATAC-seq methodology to enable the exploration of open chromatin accessibility configurations in a specific hepatocyte characteristic such as liver polyploidy were established. Specifically, polyploid murine hepatocytes, namely diploid (2n) vs. tetraploid (4n) cells may show different gene regulatory configurations according to their ploidy status.

The chromatin accessibility landscape can be assessed in single-cells by performing single-cell Assay for Transposase-Accessible Chromatin (ATAC) sequencing in individual cells (scATAC-seq) [44]. Several technical developments and methodologies have been recently established for performing scATAC-seq, among others, using microfluidics [44, 347] and liquid deposition systems [61]. Two published techniques provide distinct advantages, one being published by Buenrostro *et al.* in 2015[44] and one by Chen *et al.* in 2018 [60]. The pipeline developed by Buenrostro *et al.* 2015 consists of tagmentation per cell using the C1 Single-Cell Auto Prep System with its Open App™ program (Fluidigm, Inc.) and an improved transposase-based library preparation strategy to perform single-cell ATAC-seq [348]. This methodology relies on the distribution of the cells into microwells according to a Poisson distribution. Once distributed, a flow of buffers is created to wash and lyse the cells, transpose and tagment the DNA. Thereafter, EDTA-mediated $MgCl_2$ quenching, PCR amplification, and finally, cell harvesting are performed. After cell barcoding, library preparation and sequencing are performed in bulk. On the other hand, the protocol developed by Chen *et al.* relies on an upfront nuclei tagmentation in bulk, followed by FACS sorting the nuclei into 384-well plates containing lysis buffer composed of SDS and proteinase K together with Illumina N5 and N7 indexes. Thereafter, the Tn5 enzyme is released and the SDS is quenched with Tween-20 followed by library PCR amplification, pooling, and DNA clean-up before sequencing [60]. The two methodologies differ in several aspects, the main one being the tagmentation reaction in bulk or per single nuclei, yet the use of one or another relies on the accessibility to the necessary machinery, the affordable cost, and the needed throughput according to the research question needs and experimental approach.

Polyploidization is a characteristic feature of hepatocytes in the liver but the epigenetic landscape of polyploid cells has to date not been well studied at the single-cell level. Several transcriptomic studies have emerged to interrogate the expression profiles of the various ploidy statuses comprised in the tissue, mainly 2n and 4n [19, 42]. In order to tackle this gap, here, the combination of the

aforementioned methodologies using a plate-based approach combined with single-well tagmentation adjusting the concentration of the transposase enzyme per cell was explored.

After performing nuclei isolation from frozen mouse liver biopsies (Methods), the nuclei were FACS sorted individually into 384-well plates containing the lysis buffer composed of SDS and proteinase K together with Illumina N5 and N7 indexes [60] (Methods). The forward scatter-area (FSC-A) is used to discern the size of the nuclei, whereas the side scatter (SSC) refers to the internal complexity of cells, such as granularity. This first gating strategy was used to select all nuclei (**Figure 20**). The second plot represents the DNA content Hoechst (450/40-W vs 450/40-A), where the 2n, 4n, 8n, and 16n hepatocytes are observable (**Figure 20**B). The FSC-W vs. FSC-H as well as the SSC-W vs. SSC-H plots are used to discard doublets and purify single nuclei.

The ploidy level of the nuclei was observable in the gating strategy when loading the nuclei for FACS sorting before undergoing tagmentation in all gating settings, especially in the histogram representing the genome content staining with Hoechst (450/40-A) (**Figure 20**). The population sorted into the 384-well plates was P4 (**Figure 20**D), which comprises both 2n and 4n hepatocytes. In order to separate the different ploidy statuses, 2n and 4n nuclei were sorted according to their genome content using Hoechst. The peaks in the histogram (450-40/A vs. Counts) correspond to the DNA content of 2n hepatocytes (first peak) vs. 4n hepatocytes (second peak) (**Figure 20**E). From the total of nuclei created after doublet removal, 17.8% of hepatocytes were diploid vs. 43.2% of nuclei were tetraploid.

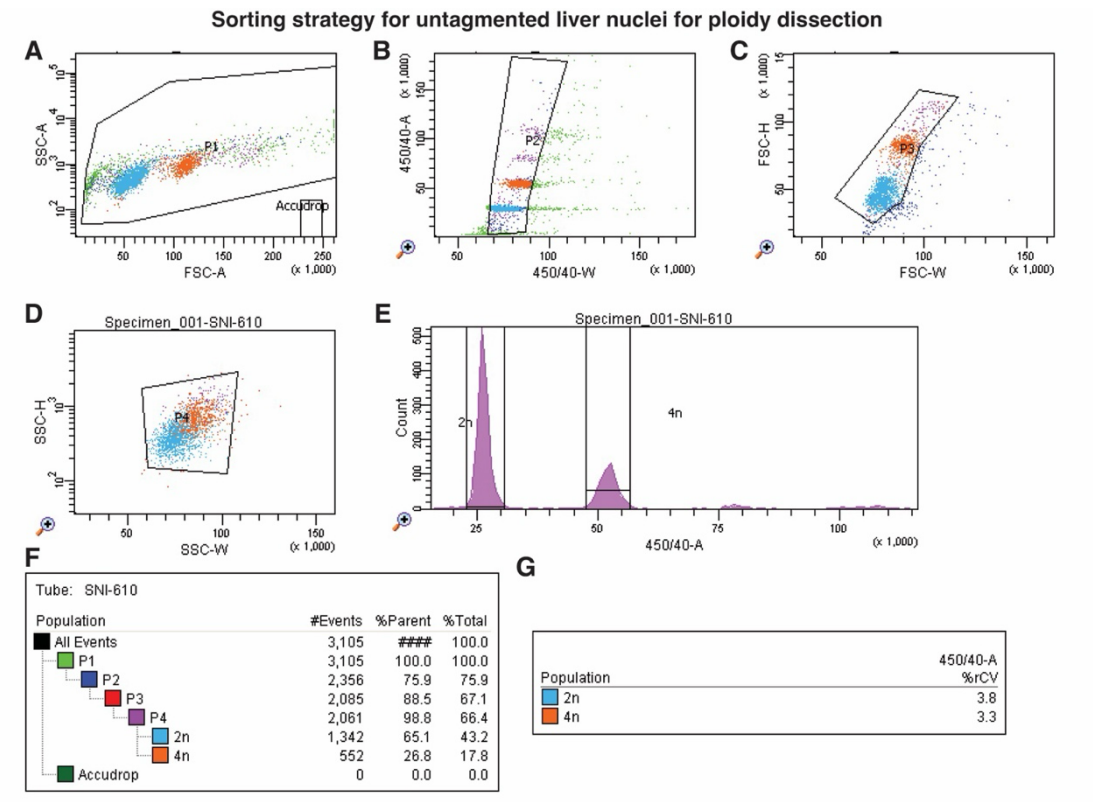


Figure 20. Nuclei isolated from flash-frozen liver biopsies can be FACS sorted according to their DNA content. ***(A)*** *Gating on forward scatter-area (FSC-A) vs. side scatter-area (SSC-A) defining P1.* ***(B)*** *Gating on 450/40 filter for violet laser width vs. area, defining P2.* ***(C)*** *Gating on forward scatter width vs. height to remove doublets and define P3.* ***(D)*** *Gating on side scatter width vs. side scatter height defining single 2n and 4n nuclei into P4.* ***(E)*** *Histogram created on the 450/40 filter vs. counts depicting peaks corresponding to the DNA content of the nuclei.* ***(F)*** *Summary table showing the number of events, the percentage of the nuclei from the parent population and the percentages of the total of nuclei.* ***(G)*** *Table showing the percentage of 2n and 4n nuclei.*

The gating strategy published in the methodology developed by Chen *et al.* [60] was recreated to sort already tagmented nuclei. However, when the tagmentation was performed upfront and then the nuclei were loaded in the sorter, the separation according to ploidy level was no longer observable along the DNA content histogram or any other lasers using P3 (450/40-A vs. FSC-H) as sorted population (**Figure 21**). This indicates that since transposition and tagmentation serve to fragment DNA in open chromatin regions, the composition of the gene dosage is no longer apparent when the DNA has been tagmented. The population sorted into 384-well plates was therefore P3, containing tagmented DNA (**Figure 21**).

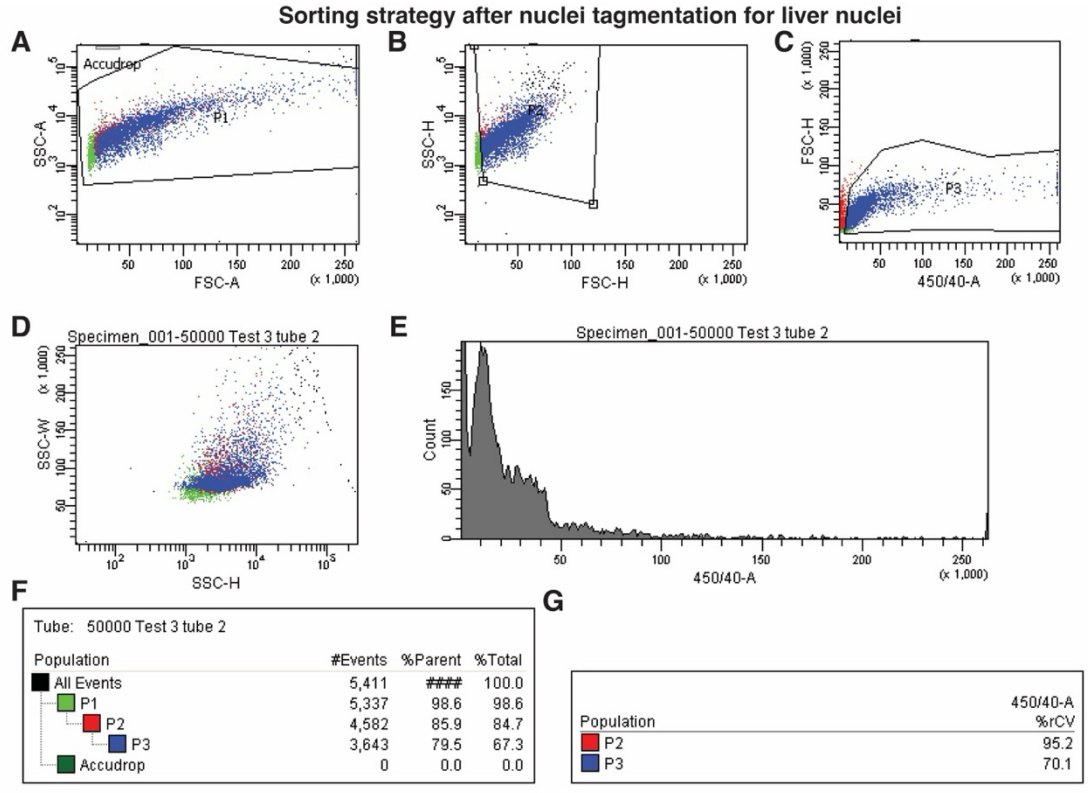


Figure 21. Nuclei isolated from flas-frozen liver biopsies can be FACS sorted according to their DNA content. ***(A)*** *Gating on forward scatter-area (FSC-A) vs. side scatter-area (SSC-A) defining P1.* ***(B)*** *Gating on forward scatter-height (FSC-H) vs. side scatter-height (SSC-H) to define P2.* ***(C)*** *Gating on the 450/40 filter-area (450/40-A) vs. forward scatter-height (FSC-H) to define P3.* ***(D)*** *Gating on side scatter width (SSC-H) vs. side scatter height (SSC-W).* ***(E)*** *Histogram created on the 450/40 filter vs. counts depicting peaks corresponding to the DNA content of the nuclei.* ***(F)*** *Summary table showing the number of events, the percentage of the nuclei from the parent population and the percentages of the total of nuclei.* ***(G)*** *Table showing the percentage of 2n and 4n nuclei.*

Following sorting, the wells containing tagged nuclei were pooled and library preparation was performed. The library assessment resulting in electropherograms showed the classical pattern of nucleosome signal periodicity of ATAC-seq libraries, with nucleosome signals peaking at around 150-200 bp [53], also coinciding with helical pitch of the DNA molecule [58, 349] and with the experimentally determined length of DNA wrapped around one nucleosome (147 bp) (**Figure 22**). The electropherogram of the untagmented DNA control sample was characterized by the absence of observable peaks (**Figure 22**). Respectively, 50, 25 and 10 cells per well were processed in samples number 1, 2 and 3 (**Figure 22**).

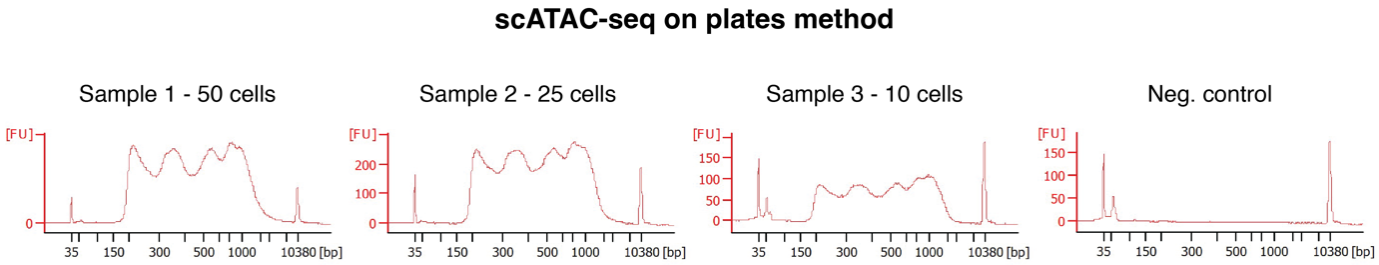


Figure 22. Representative nucleosome periodicity profiles can be detected in Bioanalyzer electropherograms of plate scATAC-seq libraries following the upfront tagmentation approach. ***(A)*** *Sample 1 containing, 50 cells per well;* ***(B)*** *Sample 2 containing 25 cells per well;* ***(C)*** *Sample 3 containing 10 cells per well;* ***(D)*** *Negative control sample containing untagmented DNA.*

The expected DNA fragmentation pattern was obtained after the scATAC-seq experiment (**Figure 22**) with a resolution down to 10 nuclei simultaneously using the protocol published by Chen *et al.* using an upfront tagmentation strategy.

Next, the combination of both strategies by sorting the untagmented nuclei to be able to study ploidy statuses and perform the nuclei tagmentation per well was performed. The electropherogram image shows the peaks for the ladder for each of the two samples (**Figure 23**A). Nevertheless, the electrophoresis profiles corresponding to 2 nuclei showed resemblance with the nucleosome periodicity, however lower fluorescent signal (**Figure 23**B). Sample 1 was obtained with the tagmentation mix published by *Chen et al.*, including Digitonin in the buffer as well as a Tagmentation Stop Buffer. On the other hand, Sample 2 was obtained using the tagmentation mix published by *Buenrostro et al*. without digitonin nor tagmentation stop buffer (Methods, **Figure 23**B).

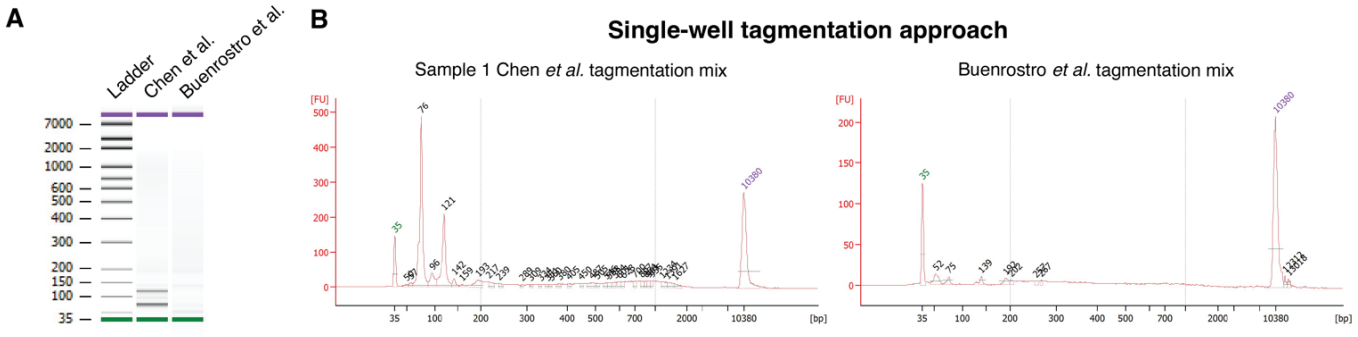


Figure 23. Representative nucleosome periodicity profiles can be detected in Bioanalyzer electropherograms of plate scATAC-seq libraries *tagmented per well.* ***(A)*** *Electrophoresis representation including ladder, sample 1 and sample 2.* ***(B)*** *Electropherograms of sample 1, obtained using the tagmentation mix published by Chen et al. and* ***(C)*** *Applying the procedure published Buenrostro et al.*

The combination of these two methodologies provides the potential to overcome the need for special equipment, expensive supplies, and a large amount of Tn5 transposase enzyme by making the reaction more efficient. Furthermore, this facilitates and eases the exploration of the chromatin accessibility landscape in complex tissues with distinctive features that require specific experimental setups and experimental optimization. An example of this is the study of polyploidy, a defining feature of hepatocytes. The ploidy levels of hepatocytes could not be dissected when the DNA tagmentation was performed prior to nuclei sorting. Therefore, this combination of scATAC-seq methodologies provides an advantage to dissect differential chromatin openness in hepatocyte polyploidy. This is achieved by combining FACS sorting of the isolated nuclei according to their ploidy status, namely their DNA content, before performing the DNA tagging and tagmentation steps in each single well containing a nucleus.

Among the characteristics of hepatocytes, polyploidy is a key feature of the liver parenchymal cells, which also perform a plethora of metabolic functions, such as the metabolism of endogenous and exogenous substances. In the next chapter of my dissertation, I will explore the metabolic function of individual hepatocytes at the epigenetic level using scATAC-seq applying a high-throughput droplet-based approach (10X Genomics).

II. Block II (Chapters III and IV): High-throughput single-cell multiomics for dissecting cellular heterogeneity in primary human hepatocytes using a droplet-based approach.

In the second block of my dissertation, the heterogeneity intrinsic in a seemingly homogeneous population of cells such as primary human hepatocytes (PHHs) *in vitro* was investigated, employing complementary high-throughput single-cell genomic approaches at the transcriptomic (scRNA-seq) and epigenomic (scATAC-seq) levels. The rationale behind the research question stated for the projects presented here was to investigate the cellular heterogeneity among PHHs exposed to a drug-related metabolic challenge, such as the coadministration of a cocktail of five drugs. In addition to that, to interrogate how chronic intracellular fat accumulation occurring in chronic diseases such as non-alcoholic fatty liver disease (NAFLD) affects the metabolic capacity of individual cells. Moreover, the response at single-cell resolution of the impact of fat accumulation on the drug-related metabolic capacity of single hepatocytes was evaluated.

Preamble: In vitro models of liver and their applicability to study liver physiology.

Method optimization

First, the experimental conditions and design selected for the single-cell experiments were previously empirically optimized and established *in vitro* using the immortalized HepG2 and HepaRG human hepatoma cell lines, which retain the main properties of primary hepatocytes [242].

The Sanofi-Aventis phenotyping cocktail composed of five drugs was used to simultaneously monitor the expression levels and the chromatin accessibility of the main five cytochrome P450 enzyme (CYP450) genes as output of the metabolic capacity of HepaRG cells (in bulk analysis) and PHHs (at single-cell level) [285, 287].

Therefore, a 66-hour incubation with this phenotyping cocktail, consisting of a mixture of individual selective substrates of CYP2D6 (metoprolol), CYP2C19 (omeprazole), CYP2C9 (S-warfarin), CYP3A4 (midazolam) and CYP1A2 (caffeine), was used to monitor changes in the *CYP450* enzymes gene expression at mRNA level [125] (Methods). All these substrates for CYP-mediated metabolism are used in index clinical drug-drug interaction studies [125]. Of these, omeprazole has a time and

dose-dependent inhibitory effect on *CYP2C19* and *CYP3A4* [153, 294], as well as working as *CYP1A2* inducer [125, 350]. Metoprolol has been shown to inhibit *CYP2D6* and *CYP3A4*, without interfering with the CYP2D6-mediated metabolism of midazolam [301]. On the other hand, caffeine, midazolam, and S-warfarin are inducers of the enzymes *CYP1A2*, *CYP3A4,* and *CYP2C9*, respectively [351-353]. Changes in CYP450 enzyme activity have been shown to be altered by lipid metabolism, showing a direct correlation [207, 225, 354]. Specifically at single-cell level, hepatic steatosis has been found to change the transcriptomic profile of parenchymal and non-parenchymal cells, as well as the cellular composition in the liver [106, 210, 211, 355].

Particularly in hepatocytes, lipid metabolism has been found to be altered upon fat accumulation due to the alteration of key enzymes in the lipid synthesis, storage, and clearance pathways [106, 356, 357]. Moreover, increased chemokine production has been observed associated with the inflammation process happening in NAFLD [358, 359]. To mimic *in vitro* the hepatic steatosis occurring in NAFLD and study its effect on liver cells at single-cell resolution, the cells (either HepG2, HepaRG, or PHHs) were loaded with intracellular lipid droplets by incubating them in FFA medium for 66 h and their individual transcriptional profile analyzed (Methods).

In vitro, benign chronic hepatic steatosis can be modeled by incubating the cells with free fatty acids (FFA) [236, 260, 360]. Diverse cellular phenotypes have been described when exposing cells to different fatty acid molecules, concentrations, combination ratios, or incubation times [260, 361, 362]. The most widely used is the combination of oleic acid (OA) and palmitic acid (PA), the two most abundant circulating fatty acids in the human body, which accumulate in plasma during human hepatic steatosis [263, 265, 363, 364]. Saturated fatty acids, such as PA, are known to produce cytotoxic effects when used alone, while unsaturated FFA like OA protects against it [365]. Intracellular lipid accumulation in human hepatocytes and hepatoma cell lines increases with increasing molar concentrations of FFA in media and incubation time [260]. It has been shown that up to a concentration of 0.25 nM of the 2:1 mixture (OA:PA) produced intracellular fat overloading after 36 h incubation, without impacting cell viability nor inducing early apoptosis [260]. To optimize the experimental conditions applied to the experiment at single-cell resolution using PHHs, the concentration and time of incubation to mimic benign hepatic steatosis *in vitro* were tested using the hepatoma cell lines HepG2 and HepaRG (Methods).

To examine the effects of the incubation with FFAs on HepG2 cells, three different concentrations (100 µM, 150 µM, and 200 µM) were tested for 72 h, using proliferation medium (DMEM+4.5% glucose) as negative control (Methods). Thereafter, 4% paraformaldehyde fixation and Oil Red O

staining of intracellular lipids were performed and assessed under brightfield microscopical observation (Methods, **Figure 24**).

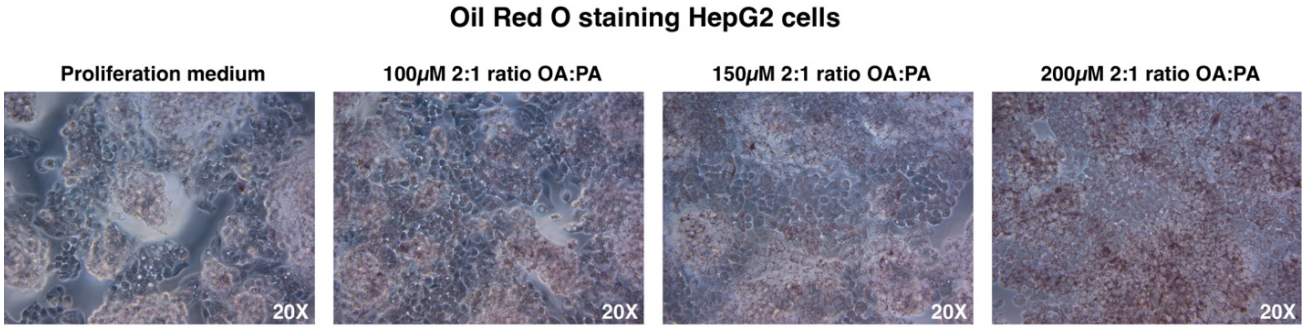


Figure 24. Increasing concentration of oleic to palmitic fatty acids leads to increased intracellular fat accumulation in HepG2 cells. *Microscopical images of Oil Red O stained HepG2 hepatoma cell line cells in three different concentrations of the FFA mixture composed of oleic acid (OA) and palmitic acid (PA) for 72 h.* ***(A)*** *Control sample in proliferation medium.;* ***(B)*** *Sample incubated in 100 µM ratio OA:PA.* ***(C)*** *Sample incubated in 150 µM ratio of OA:PA.* ***(D)*** *Sample incubated in 200 µM ratio OA:PA.*

Cells incubated in 200 µM of a 2:1 ratio of OA to PA showed the highest accumulation of Oil-red-O stained intracellular lipids compared to the two other lower concentrations of the FFA mixture and to the negative control with no additional cell detachment, cell death, or clonal expansion arrest (**Figure 24**). Therefore, this concentration and ratio of FFA were selected for the subsequent experiments in HepaRG cells (in bulk) and PHHs (at single-cell resolution).

While HepG2 cells are a valuable tool for *in vitro* steatosis studies [260], they have poor drug-metabolizing competence, as they lack functional expression and inducibility of nearly all relevant cytochrome P450 enzymes [229, 366]. Therefore, other liver *in vitro* models such as the metabolically competent HepaRG cell line were developed [236, 367]. This hepatoma cell line conserves the expression of several P450 enzymes, phase II metabolism-related enzymes, and nuclear receptors such as CAR and PXR at levels comparable to those of PHHs [236]. Thus, they are more suitable and closer model to explore drug-related induction studies or enzymatic activity analysis to assess metabolic response to chemical stimulus or for the assessment of drug-induced liver injury (DILI) [229, 236, 368]. Morphologically, HepaRG cells resemble the *in vitro* structures found in PHHs, forming trabeculae on adherent culture surfaces after 72 h of incubation (**Figure 25**).

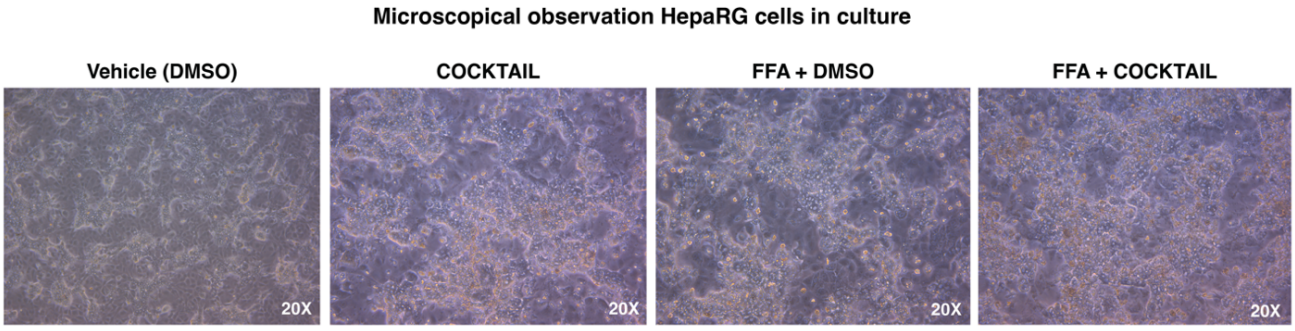


Figure 25. Microscopical observation in the brightfield of HepaRG cells in the four experimental conditions after 72 h incubation. *(**A**) Vehicle (DMSO) treatment. (**B**) Five-drug cocktail (Cocktail) treatment. (**C**) Free Fatty Acids+Vehicle (FFA+DMSO) and (**D**) Free Fatty Acids+Five-drug cocktail (FFA+Cocktail) treatment.*

HepaRG cells represent a closer model to PHHs, and therefore the experimental conditions designed for the exploration of PHHs at single-cell resolution were first optimized in HepaRG cells in bulk (Methods). Hence, cryopreserved HepaRG cells were plated and incubated for a total of 72 h (Methods). Of those, a 66-h incubation with either vehicle (DMSO) or the Sanofi-Aventis phenotyping five-drug cocktail (metoprolol, omeprazole, S-warfarin, midazolam, and caffeine) was used to examine changes in the *CYP450* enzymes expression in bulk by qPCR [125] (Methods). Hereby the functionality of the cocktail to induce the expression of the main five CYPs responsible for the metabolism of the drugs was tested. Chronic fat accumulation was achieved by incubating the cells with a 200 µM FFA mixture of a 2:1 ratio of oleic acid to palmitic acid for a total of 72 hours. Four conditions were studied: a) Vehicle (DMSO 0.5% v/v); b) Cocktail (66h incubation with a five-drug cocktail); c) FFA+DMSO (2:1 ratio oleic:palmitic free fatty acids) and d) FFA+Cocktail (FFA incubation + five-drug cocktail).

After 72 h incubation, cells were collected, and total RNA was isolated, followed by cDNA synthesis and qPCR (in technical duplicates) (Methods). The expression level was calculated using the fluorescence data (Ct values), and the fold change over the housekeeping gene *PGBD* was calculated using the methodology published by Ramarkes *et al.* 2003 [311] taking into consideration the efficiency of each primer for its target (Methods). The efficiency calculated for the four targets was the following: 1.58 for *PGBD*, 2.34 for *CYP2C9*, 2.02 for *CYP2C19,* and 3.22 for *CYP3A4*. This is higher than 2, which can influence the fold change analysis and could be due to overconcentrated samples [369, 370]. Then, the obtained fold change (target vs. PGBD) values were $\log_{10}$-transformed, and the mean and standard deviation of the two technical replicates were calculated. The expression level under DMSO treatment was set to baseline and the $\log_{10}$-fold change per each target (*CYP2C9,*

CYP2C19, CYP3A4) per treatment condition (Cocktail, FFA+DMSO, FFA+Cocktail) was calculated (**Figure 26**).

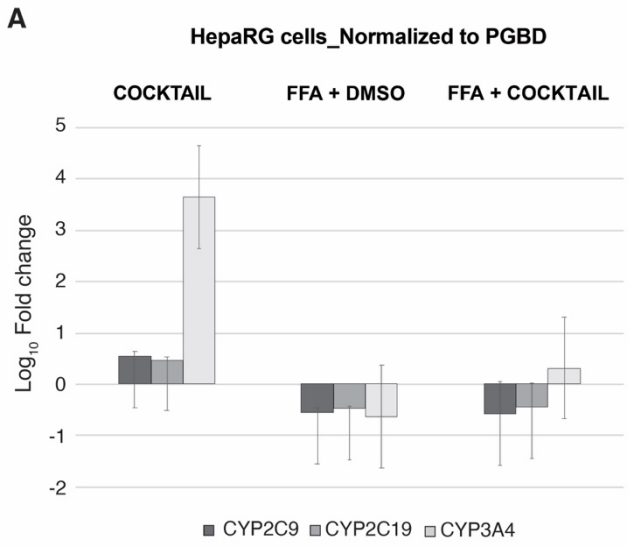


Figure 26. (A) Cocktail treatment of HepaRG led to the induction of the three CYPs, whereas fat accumulation impaired the metabolism of the drugs. *HepaRG cells treated with Vehicle (DMSO), Cocktail (five-drug Sanofi-Aventis cocktail), FFA+DMSO, or FFA+Cocktail. mRNA expression levels normalized to the housekeeping gene PGBD. DMSO expression levels were set to baseline. Error bars: Mean+SD.*

Incubation with the five-drug cocktail led to the induction of the mRNA levels of *CYP2C9*, *CYP2C19,* and *CYP3A4* showing a 0.55-fold, 0.48-fold, and 3.65-fold induction over DMSO baseline level, respectively (**Figure 26**). A downregulation in the mRNA levels of *CYP2C9, CYP2C19,* and C*YP3A4* compared to DMSO baseline levels was observed upon free-fatty acid (FFA+DMSO) treatment leading to intracellular lipid accumulation in HepaRG cells. The largest downregulation was observed for CYP3A4, showing a 0.62-fold.

When cells were treated with both free fatty acids and the five-drug cocktail (FFA+Cocktail), an upregulation of the mRNA level of *CYP3A4* was measured, compared to FFA+DMSO treatment (**Figure 26**), reaching a 0.32-fold upregulation above the DMSO baseline level. However, no upregulation was observed for the other two targeted CYPs (*CYP2C9* and *CYP2C19*).

This indicated that the treatment with a five-drug cocktail containing substrates of five targeted CYP50 enzymes led to the upregulation of three targeted CYPs by the drugs. Furthermore, fat accumulation reduced the expression of the targeted CYP450 enzymes detected, reducing the capacity to metabolize the drugs in the five-drug cocktail of HepaRG cells. Upon the simultaneous incubation with FFA and the five-drug cocktail, only *CYP3A4* was upregulated compared, indicating that intracellular fat accumulation in HepaRG cells *in vitro* impacts the capacity of the cells to metabolize several drugs simultaneously.

CHAPTER III: Exploration of the chromatin accessibility using scATAC-seq reveals distinct chromatin accessibility configurations governing gene expression upon drug challenge.

Using the drug cocktail and the FFA incubation conditions optimized on hepatoma cell lines in the preamble, in this chapter of my dissertation, the investigation of the chromatin accessibility landscape at single-cell resolution of PHHs *in vitro* incubated with four experimental conditions using a droplet-based approach (10X Genomics) was performed. At single-cell level, scATAC-seq allows the exploration of the heterogeneity in the chromatin accessibility arrangements orchestrating the distinct and gene expression profiles observed upon different conditions, health statuses [44-50, 371]. In human liver, scATAC-seq analysis of healthy, steatotic, or fibrosis NASH has shown differentially accessible region profiles with stage-specific DNA regulatory elements allowing NAFLD subtypes [217, 372].

Primary hepatocytes are the gold standard liver *in vitro* model for the investigation of drug-related metabolism and responses in humans, and they are considered a seemingly homogeneous population of cells [229, 238, 271, 373, 374]. However, it is known that cellular heterogeneity can be found in the responses of individual cells at the epigenomic level [215-217]. Commercially available cryopreserved PHHs from four donors were used as a liver *in vitro* model to investigate the chromatin openness landscape of individual cells to a drug-related metabolic challenge and chronic fat accumulation. The PHHs from those four donors were plated and incubated for a total of 72 h, of which 66 h with either vehicle (DMSO) or a five-drug cocktail. Chronic fat accumulation, as occurring in hepatic steatosis during NAFLD, was achieved by incubating PHHs with a 200 µM FFA mixture of a 2:1 ratio of oleic acid to palmitic acid for 72 hours. A total of four conditions were studied: a) Vehicle (DMSO 0.5% v/v); b) Cocktail (66h incubation with a five-drug cocktail); c) FFA+DMSO (2:1 ratio oleic:palmitic free fatty acids), and d) FFA+Cocktail (FFA incubation plus five-drug cocktail).

After a total of 72 h incubation, 66 of them in DMSO or Cocktail, brightfield microscopical observation revealed all the conditions were phenotypically apparently homogeneous (**Figure 27**). Intracellular lipid accumulation was observable in the samples incubated with FFA, namely FFA and FFA+Cocktail samples.

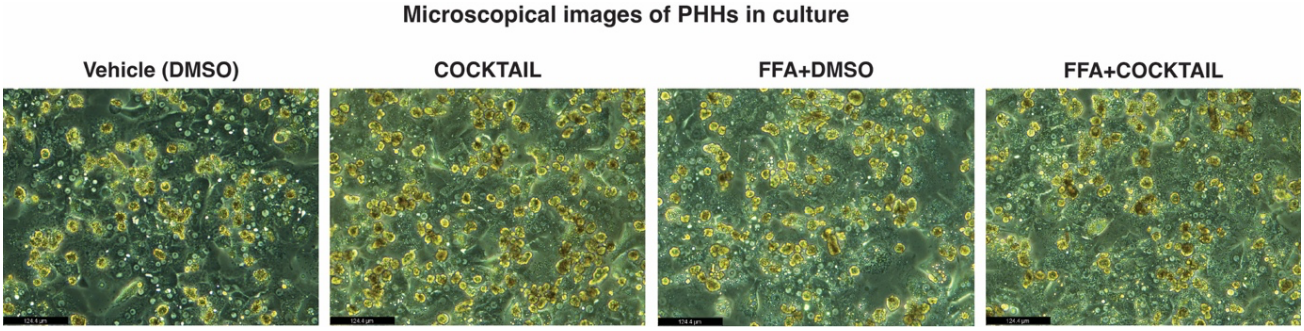


Figure 27. Brightfield microscopical images of PHH cells in the four experimental conditions after 72 h incubation. *(A) Vehicle (DMSO) treatment.* ***(B)*** *Five-drug cocktail (Cocktail) treatment.* ***(C)*** *Free Fatty Acid+Vehicle (FFA+DMSO) and* ***(D)*** *Free Fatty Acid+Five-drug cocktail (FFA+Cocktail) treatment.*

Next, to explore the heterogeneity among PHHs regarding their chromatin accessibility configuration upon cocktail treatment, intracellular fat accumulation, or both simultaneously, scATAC-seq was performed using the 10X Genomics Chromium platform (Methods).

The experimental approach for performing scATAC-seq includes a nuclei isolation step followed by nuclei lysis, tagmentation and transposition, library preparation, and sequencing (Methods) [53, 348]. Among these, nuclei isolation and lysis are essential steps subject to optimization in order to obtain a clean single-nuclei suspension, an efficient nuclei lysis and tagmentation reaction. Therefore, the nuclei isolation methodology published by 10X Genomics recommended optimizing the lysis time to the tissue or cell type of interest [68]. Due to known physical properties of PHHs such as a stiff cellular membrane and large cellular size, the lysis time was set for a short lysis of 10 min and a longer lysis time of 30 min. The quality control of the libraries obtained after bulk nuclei lysis, the Bioanalyzer 2100 electropherograms, and tagmentation are shown in **Figure 28**:

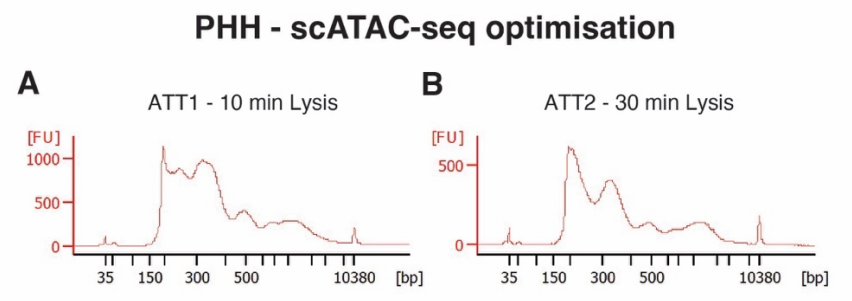


Figure 28. Comparable nucleosome periodicity profiles are obtained from two PHH samples processed for bulk ATAC-seq incubated in Lysis Buffer for 10 and 30 min, respectively. *Bioanalyzer electropherogram traces of* ***(A)*** *Sample ATT1, incubated 10 min in Lysis Buffer.* ***(B)*** *Sample ATT2, incubated 30 min in Lysis Buffer.*

The nucleosome periodicity profile obtained in both samples was comparable, with the classical profile expected from a scATAC-seq library (**Figure 28**). After sequencing, 2,945 cells were recovered for sample ATT1 vs. 7,899 cells from sample ATT2. However, the loaded number of cells was to recover

5,000 cells, hence concluding that the number of cells recovered from sample 2 was overestimated. To overcome this, the number of recovered cells was forced to 2,453 based on the barcodes in fragments overlapping peaks plot (Web summary output from 10X Genomics).

The TSS enrichment and fragment length distribution calculated showed less noise vs. signal ratio in the sample lysed for 10 minutes (**Figure 29**). Based on the QC results, the sample lysed for 10 min yielded a better quality overall, and hence, 10 min incubation in lysis buffer was selected for the experimental design.

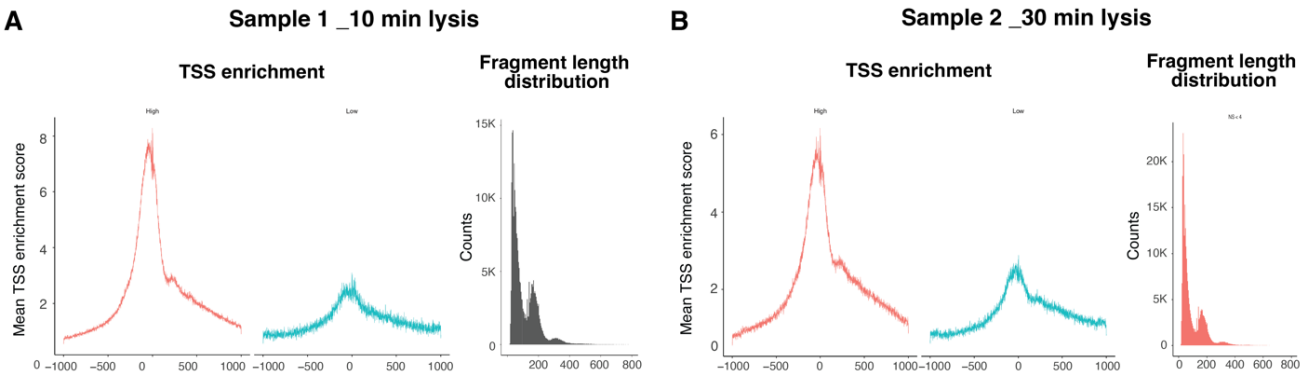


Figure 29. Comparable TSS enrichment and fragment length distribution was achieved in PHH samples incubated for 10 vs. 30 minutes in lysis buffer. ***(A)*** *TSS enrichment and fragment length distribution plots from PHHs tagmented for 10 minutes in the Lysis Buffer from 10X Genomics nuclei isolation protocol.* ***(B)*** *TSS enrichment and fragment length distribution plots from PHHs tagmented for 30 minutes in the Lysis Buffer from 10X Genomics nuclei isolation protocol. Figure computed by Mr. Patrick Hanel using Cell Ranger-called peaks.*

Microscopical inspection of the single nuclei suspension revealed that the cell lysis and isolation of nuclei was not totally efficient. Efficient cell lysis for tagmentation involves the disruption of the cellular membrane without disturbing the nuclear membrane. Since some cells showed nuclei could be observed, I sought to further optimize the cell lysis and nuclei isolation procedure for PHHs *in vitro* (**Figure 30**).

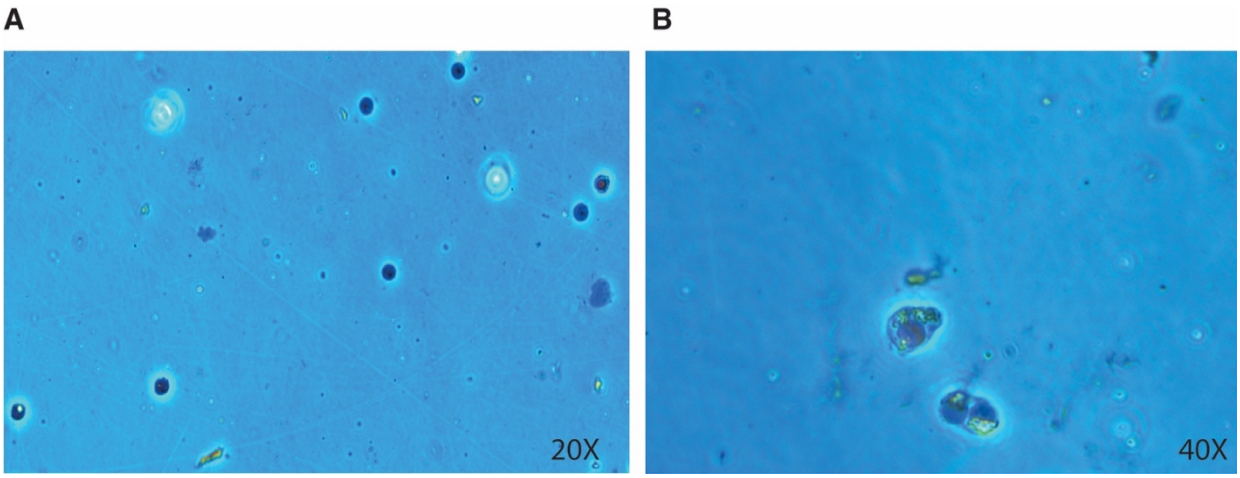


Figure 30. Microscopical image of nuclei isolated from PHHs and stained with trypan blue. *(A) Single nuclei in suspension dissociated and released from the cell membrane.* ***(B)*** *Nuclei contained in the cell membrane of PHHs indicate that the lysis was not 100% efficient.*

While the cell lysis, nuclei isolation and tagmentation optimization experiments were performed for PHHs, the first work on scATAC-seq from mouse frozen liver using a droplet-based approach provided by 10X Genomics was published by Chen *et al.* 2020 [48]. In their protocol, incubation of the cells in a hypotonic swelling buffer was used to enlarge the hepatocyte nuclei and facilitate the membrane disruption. Therefore, to test the effectivity of this swelling buffer on PHHs cultured *in vitro*, PHHs were isolated in homogenization buffer in combination with swelling buffer (Methods).

At first, the methodology was validated using murine frozen liver as the same starting material. Nuclei were isolated following the procedure by Chen *et al.* 2020 [48] minutely and lysing the nuclei for 10 min using the 10X Genomics lysis buffer. Initially, 50,000 nuclei were counted and used for tagmentation for 60 min at 1000 rpm. After library preparation, the profile obtained in the electropherogram in the Bioanalyzer traces showed the expected nucleosome periodicity (**Figure 31**A). The same experimental conditions used for mouse frozen tissue were applied to mouse fresh liver in a new experiment. The electropherogram showed nucleosome periodicity similar to the one obtained from frozen mouse liver (**Figure 31**B). In contrast, the physical characteristics of PHHs, such as a robust cellular membrane hamper the lysis of the nuclear membrane and DNA tagmentation. Therefore, when isolating nuclei from primary hepatocytes, both swelling buffer and homogenization buffer with 0.3% Triton X and 0.2% NP-40 were used (Methods). The Bioanalyzer traces showed the expected nucleosome periodicity of an ATAC-seq library and a similar profile to that obtained using frozen mouse liver (**Figure 31**C).

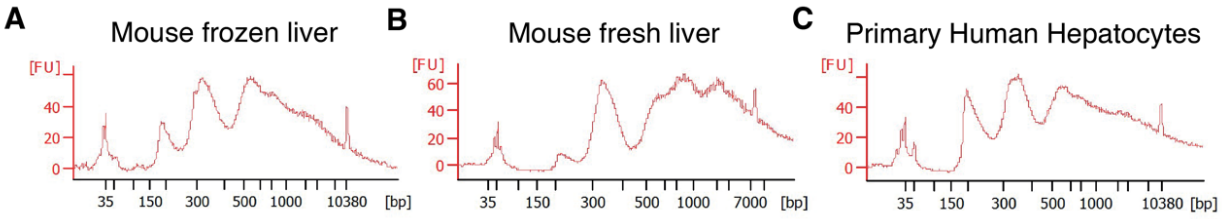


Figure 31. Similar nucleosome periodicity profiles were obtained using different sample origins. *Bioanalyzer electropherograms of bulk ATAC-seq experiments performed from* ***(A)*** *a frozen mouse liver biopsy.* ***(B)*** *a fresh mouse liver biopsy; and* ***(C)*** *primary human hepatocytes.*

To explore the chromatin accessibility configuration of PHHs in response to a metabolic challenge consisting of the exposure to a cocktail of five drugs, upon chronic fat accumulation or both of these conditions simultaneously, PHHs from two donors were incubated for a total of 72 h, 66 of them in

DMSO or Cocktail, resulting in a total of four experimental conditions: DMSO, Cocktail, FFA, and FFA+Cocktail. A total of eight samples were used to perform high-throughput single-cell ATAC-seq using a droplet-based approach (10X Genomics) (Methods) [375].

After plating, cells were microscopically inspected and differences in cellular adherence and viability as well as survival were observed between donors, affecting the subsequent total number of cells detached after incubation time (72 h). Thereafter, cells were collected to perform nuclei isolation, nucleic acid transposition, and loading in the 10X Genomics Chip (Methods). Differences in the number of cells recovered were observed due to the uneven concentration of the eight samples and the loaded number of cells into the 10X Genomics chip (**Table 6**).

Table 6. Summary of donors, samples, number of cells collected, and cells recovered in the scATAC-seq experiment (10X Genomics Chromium).

Donor	Sample	Experimental condition	No. of cells	Targeted no. of recovered cells	No. of cells recovered after Chromium
1	1	DMSO	1,000,000	1000	1,602
1	2	Cocktail	2,075,000	1000	1,088
1	3	FFA	1,100,000	1000	6,111
1	4	FFA+Cocktail	900,000	1000	1,841
2	5	DMSO	1,285,000	1000	3,441
2	6	Cocktail	1,050,000	1000	666
2	7	FFA	950,000	1000	243
2	8	FFA+Cocktail	790,000	1000	502

Nevertheless, after GEM recovery, no wetting failures were observed across the eight samples. The generated libraries were assessed in the Bioanalyzer 2100, however, the profiles obtained did not accurately resemble the representative traces provided by 10X Genomics for scATAC-seq libraries, in samples 1, 2, 3 and 8 (**Figure 32**). The nucleosome periodicity pattern is visible in the nucleosome-free and one nucleosome peaks, no longer observable for higher sizes corresponding to multiple nucleosomes, except for samples 4 and 6 where they are noticeable (**Figure 32**).

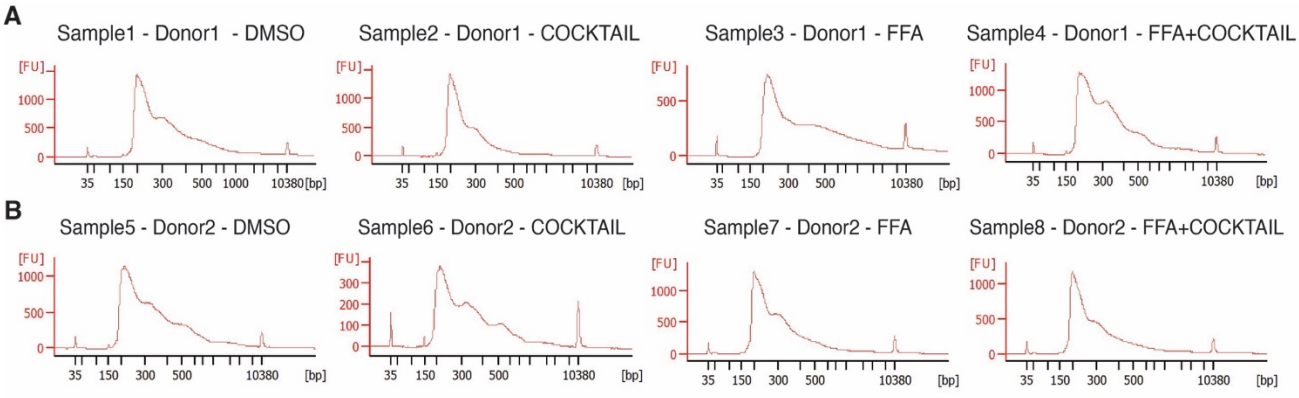
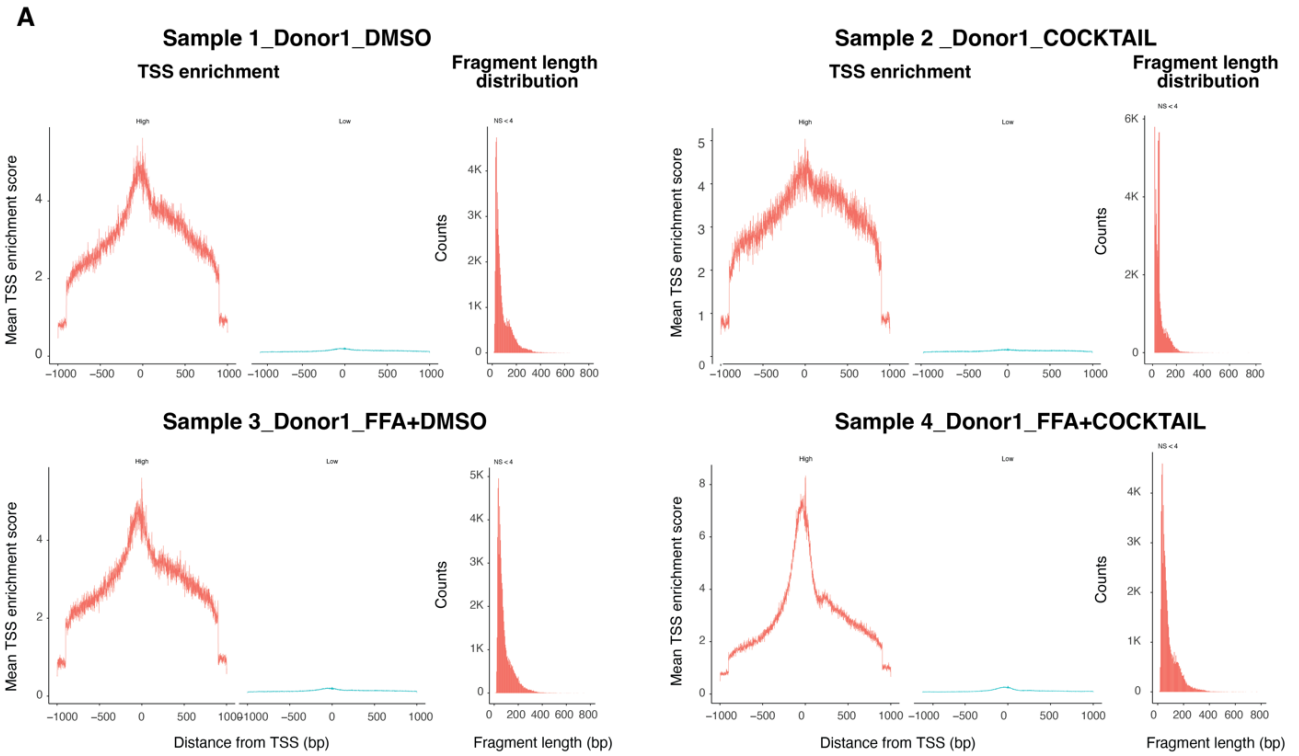


Figure 32. Nucleosome periodicity patterns show noticeable differences in the eight samples in the scATAC-seq experiment. *Bioanalyzer electropherogram traces of the libraries obtained after the scATAC-seq experiment. Libraries from the eight samples in the four treatment conditions (DMSO, Cocktail, FFA, and FFA+Cocktail) from* ***(A)*** *donor 1 and* ***(B)*** *donor 2.*

The concentration of sample 6, corresponding to Cocktail-treated cells from Donor 2, was the lowest across all samples with 4,388 pg/µL, corresponding to 666 cells recovered. However, sample 7, corresponding to FFA-treated cells was the sample with the least number of cells recovered (243) (**Table 6**). At least three times higher concentrations were obtained, being the highest for sample 1 with 13,251.95 pg/µL. The nucleosome-free signal was high in all samples, which could be an indicator of DNA tagmentation excess. Nevertheless, sample 6 showed the clearest peaks in the nucleosome distribution profile. After sequencing, the *Cell Ranger-atac* fragments file output was used to call peaks using MACS2 [322]. Peaks within 5 kb upstream of a gene body were kept to account for potential enhancers or promoters upstream of the TSS in the analysis [375].



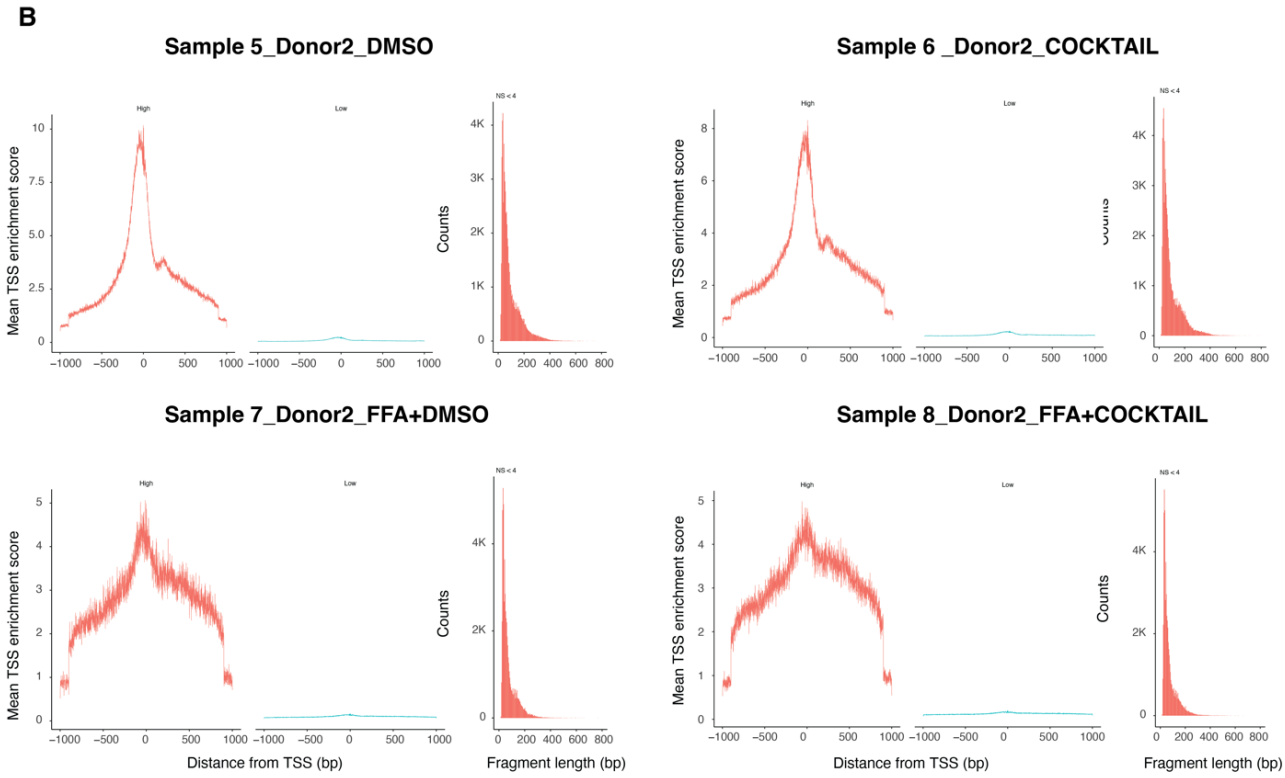


Figure 33. Different TSS enrichment and fragment length distribution was achieved for the eight samples in the scATAC-seq experiment using PHH. *TSS enrichment and fragment size distribution histograms calculated per sample using MACS2-called peaks using Signac.* ***(A)*** *Four samples corresponding to the four treatment conditions from donor 1.* ***(B)*** *Four samples corresponding to the four treatment conditions from donor 2.*

The peaks called individually per sample were used to calculate the TSS enrichment and the fragment length distribution for quality control using *Signac* [376] (**Figure 33**).

Thereafter, the generated peaks were combined using *Bedops* [323] and a count matrix built based on peaks using *EpiScanpy* [324](Methods). Cells with less than 100 peaks were filtered out and a minimum of peaks present in at least 10 cells were kept. Filtering thresholds were set by observing the histograms showing the number of open features per cell and number of common features among cells (**Figure 34**).

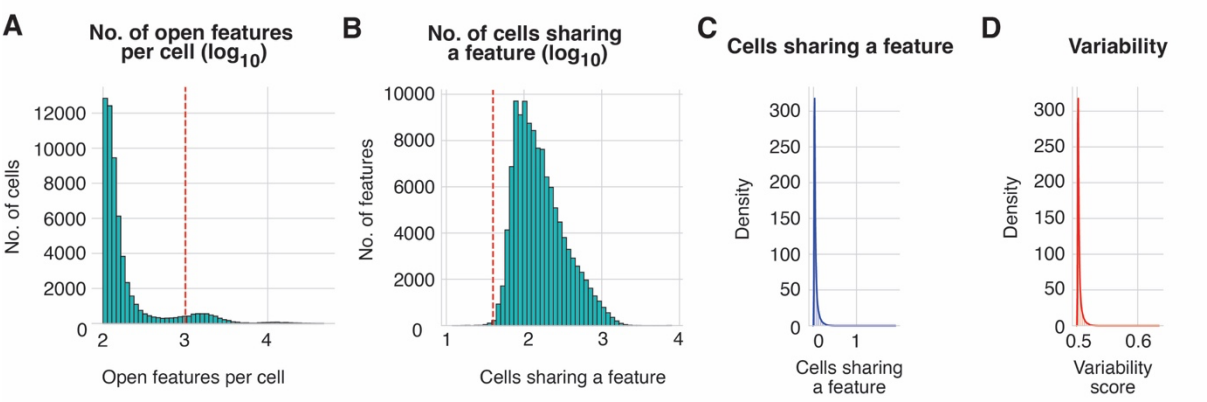


Figure 34. Filtering criteria applied to the scATAC-seq dataset. *(A) Histogram showing the number of open features per cell (log_{10}).* ***(B)*** *Histogram depicting the number of cells sharing a feature (log_{10}).* ***(C)*** *Density plot showing the cells sharing a feature.* ***(D)*** *Density plot showing the variability score of all cells.*

Differences in coverage and feature recovery were detected due to the uneven cell recovery of the loaded cells in the chip across the eight samples. After filtering, the 50,000 highly variable peaks were selected, resulting in the removal of the first sample of eight from the analysis (**Figure 35**A). This sample corresponded to DMSO-treated cells from donor 1, which had a total of 1,601 recovered cells, however only a total of 2,633 peaks were detected in this sample (Table 2, **Figure 35**A), representing the least covered sample.

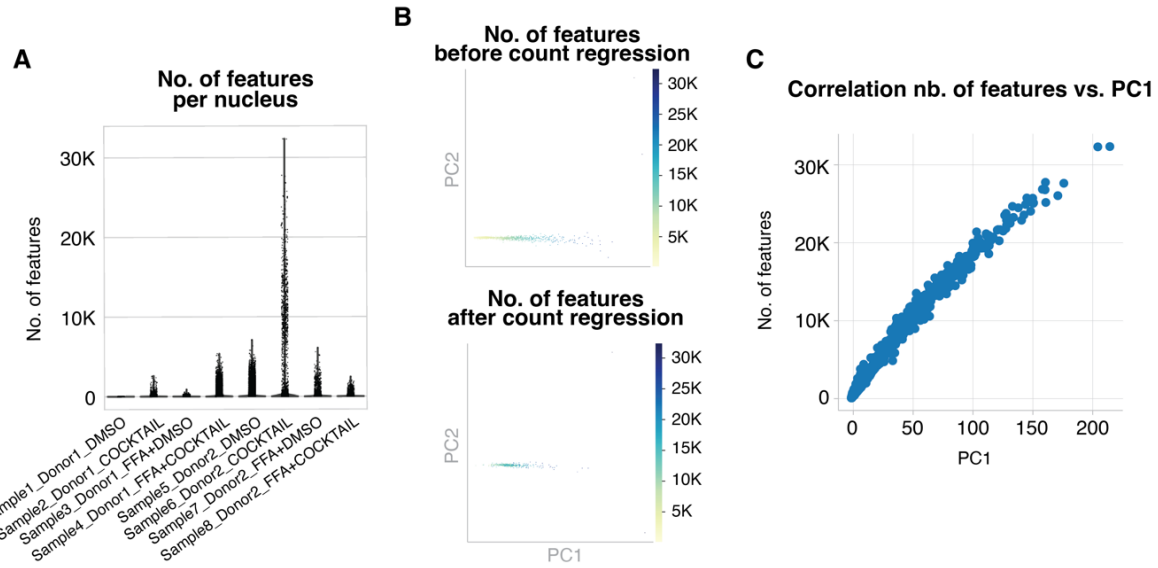


Figure 35. Differences in the number of features and coverage were identified among the eight samples. *(A) Violin plot depicting the number of features per nucleus in each of the eight samples;* ***(B)*** *PCA showing the number of features along PC1 before and after regressing the number of counts. Color scale corresponds to the number of features. K=1,000.* ***(C)*** Correlation between the number of features and principal component 1. Correlation=0.99; p value=0.0.

Additionally, the number of features detected per nucleus was highest for the sample corresponding to the Cocktail treatment sample from donor 2 (**Figure 35**A). This sample had a low number of cells

recovered (666, **Table 6**), and the highest coverage across all samples with a total of 11,849,497 features detected.

The number of features in differently covered samples can be a confounding factor for differential chromatin openness analysis. To validate whether the coverage was a confounding factor, PCA was calculated on the number of features before and after regressing the number of counts, showing that they spread along PC1 and hence indicating that indeed they were driving variation (**Figure 35**B). In order to validate this, the correlation of those with PC1 was calculated, resulting in R^2=0.99 (p value=0.0) (**Figure 35**C). Based on that, the number of counts was regressed out as a normalization strategy "*epi.sc.regress_out"* function of *episcanpy* (Methods). The UMAP representation of the clustering of the cells per sample, Louvain cluster, number of features, donor and treatment are shown in **Figure 36**. The overall data structure appeared clearer after regressing out the number of features, yet a severe change was not observable.

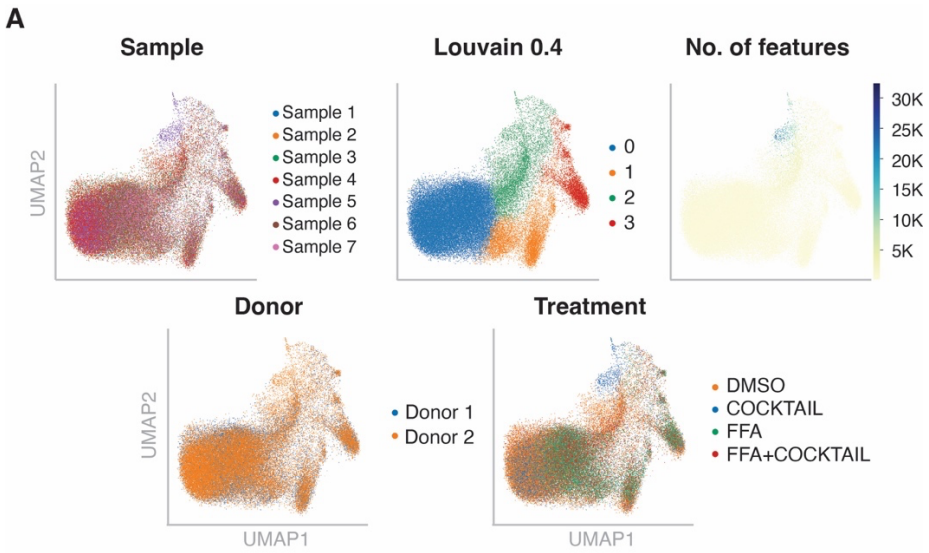


Figure 36. Homogeneous distribution of cells across donors and treatment was identified, although a higher number of features were detected in one of the eight samples. (A) *UMAP colored by the clustering per sample, Louvain clusters (resolution = 0.4), number of features, donor, and treatment (DMSO, Cocktail, FFA, and FFA+Cocktail) before number of counts regression.*

Doublet scoring was performed by using *Scrublet*, defining the expected overall doublet rate at 6.2%, as described by 10X Genomics in the scATAC-seq v1.1 User Guide [321], and no doublets were detected in the dataset.

Thereafter, to characterize the groups by ranking differentially open peaks between the four conditions in the study (DMSO, Cocktail, FFA, and FFA+Cocktail), the *scanpy* function "*sc.tl.rank_genes_groups"* was used (**Figure 37**). Genes 5 kb downstream of open peaks were investigated. The peaks were annotated by genes located 5 kb downstream of the peak to study gene

regulation. This assumes that chromatin accessibility at gene regulatory regions across the genome is an indicator of gene expression and affects the transcriptional outcome [377, 378].

Upon Cocktail treatment, the differentially open peaks corresponded mostly to intergenic regions. These constitute the majority of the genome and are non-protein-coding, typically enhancers and promoters [379, 380], known to regulate the expression of adjacent genes. Among the top 20 peaks in FFA+Cocktail-treated cells, a peak upstream of *CEACAM1* was detected, which is a transmembrane glycoprotein involved in differentiation and polarity maintenance and is expressed during hepatocyte differentiation and liver regeneration [381]. Additionally, oxidative stress-related genes such as *SOD2* and *LIMD1/LARS2* were also detected downstream of differentially open peaks (**Figure 37**).

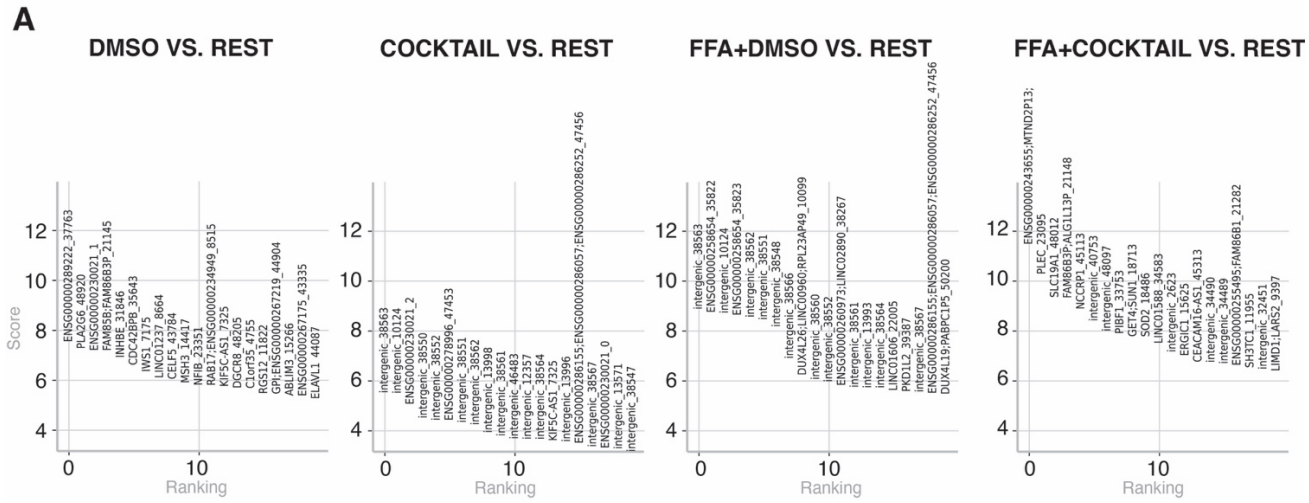


Figure 37. Ranking of genes in the four experimental conditions. *(A) Gene ranking defining the differentially accessible regions in the four treatment conditions compared to the other three (DMSO, Cocktail, FFA, FFA+Cocktail).*

Next, to assess the chromatin accessibility configurations under the four treatments, differential peak accessibility analysis was performed (**Figure 38**). Upon DMSO treatment, peaks in proximity to cellular maintenance-related genes were detected, such as the cell trafficking protein *NAPA* also known as α-*SNAP* [382], the adaptive response gene *ATF3* [383, 384], and the serum protein albumin *(ALB),* the most abundantly secreted protein by hepatocytes [72].

Upon Cocktail treatment, peaks upstream of the phase I metabolism genes such as the cytochromes were accessible for transcription. For instance, *CYP2C9*, *CYP3A5*, *CYP2B6,* or *CYP1A2,* were not detected in DMSO-treated cells (**Figure 38**). Hence, exposure to a five-drug cocktail induces a

chromatin accessibility change that is specific to exposure to a cocktail of drugs in PHHs *in vitro* compared to vehicle treatment.

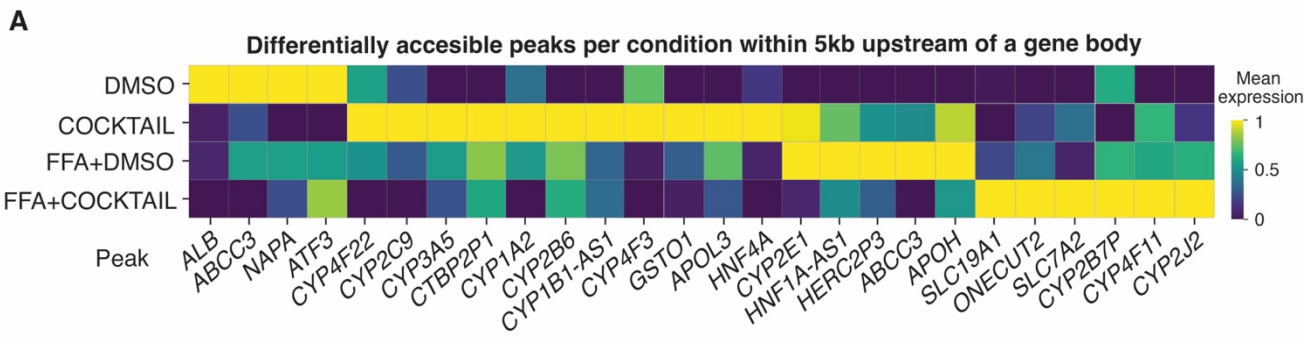


Figure 38. Differential peak accessibility of endo- and exogenous genes was detected between treatment conditions. ***(A)*** *Matrix plot depicting the accessibility of genes belonging to the 200 most highly variable features in DMSO, Cocktail, FFA and FFA+Cocktail-treated cells. Color scale represents the mean expression.*

Upon FFA treatment, regions upstream of specific lipid metabolism-related genes were differentially open. For instance, the multidrug resistance protein and phase III transporter *ABCC3* [385], the apolipoprotein transporter *APOH* associated with diabetes and metabolic syndrome [386], and the lipid metabolism-related cytochrome *CYP2E1* associated with the development of NAFLD and NASH progression [387] (**Figure 38**). It has been reported that deletion of HNF1A in mouse hepatocytes leads to fatty liver-related hepatocellular carcinoma [388] and in humans, that it can ameliorate NAFLD and NASH [389]. This hints that fat accumulation leads to a distinct chromatin openness configuration involving genes related to lipid metabolism and accumulation in hepatocytes.

Lastly, in primary hepatocytes treated with FFA+Cocktail, regions upstream of *CYP450* enzyme genes related to the metabolism of sterols such as *CYP2J2* and eicosanoids *CYP4F11*, which plays a crucial role in inflammatory processes [390] were differentially accessible. In addition, regions in proximity to solute carriers were identified as differentially accessible in FFA+Cocktail-treated cells. For instance, the folate transporter *SLC19A1* [391, 392] which has been reported to be involved in lipid droplet accumulation in hepatocytes [393]; and the amino acid transporter *SLC7A2*, which has been previously related to inflammation and immunity diseases such as HCC [394, 395] (**Figure 38**). The transcription factor *ONECUT2* has been reported to have reduced motif activity in advanced NASH disease [396]. This indicates that, upon exposure to both intracellular fat accumulation and a five-drug cocktail, the chromatin accessibility arrangement of PHHs changes to open regions in proximity to genes related to inflammation and cellular transmembrane transport.

Taken together, this suggested that the experimental methodology here developed and used in PHHs in vitro for chromatin accessibility assessment allows the preliminary identification of differential open regions in response to a metabolic challenge such as exposure to a cocktail of drugs, chronic intracellular lipid accumulation, or both conditions simultaneously.

CHAPTER IV: Transcriptomics analysis of primary human hepatocytes shows differential metabolic responses.

Analysis of factors affecting inter-donor variability and liver drug-related metabolic capacity.

In the previous chapter (Chapter III), a genomic feature such as chromatin accessibility differences in primary human hepatocytes in response to a five-drug cocktail treatment, intracellular fat accumulation, or both simultaneously were interrogated. Therefore, four experimental conditions were analyzed: Vehicle (DMSO), five-drug cocktail (Cocktail), free-fatty acids (FFA), and the combination of free fatty acids + five-drug cocktail (FFA+Cocktail) (Methods).

As previously mentioned, the current gold standard liver *in vitro* model for the investigation of drug-related metabolism and responses are PHHs, which are considered a homogeneous cell population [229, 237, 238, 271, 373, 374]. Assuming homogeneity in the cellular responses towards treatment, the assessment of efficacy, safety, and toxicity is performed in bulk analyses in the early phases of drug discovery and development using PHHs. However, scRNA-seq has revealed the presence of cellular heterogeneity in the liver in both healthy and diseased statuses [2, 5, 6, 84, 210, 397-401]. Cellular heterogeneity in the liver is affected and driven by several factors, among which is the presence of fat accumulation in the form of triglycerides in hepatocytes [106, 210], known as steatosis, directly related to the development of NAFLD and associated with metabolic dysfunction, inflammation, and increased risk for the development of fibrosis [402-405]. Other factors impacting the heterogeneity found in the transcriptomic level is aging, which has been shown to increase transcriptional noise in several tissues and organisms [185, 186, 188, 406]. In addition to these, the presence of Single-nucleotide polymorphisms (SNPs) increases the inter-individual variability that has been observed in the metabolism of CYP substrates *in vivo* [407, 408]. Across the genome, the presence of mutations in the nucleotide sequence in gene bodies (coding sequences) can lead to an alteration in the amino acid sequence configuring protein after translation and modification of the protein activity (known as non-synonymous SNP or mutation) [409, 410]. In the case of the CYP450 superfamily of enzymes, it is known that the presence of SNPs in the gene bodies can affect the metabolic capacity of the individuals, conferring different phenotypes among the population based on their capability to metabolize substances [117, 411].

Single-nucleotide polymorphisms introduce interindividual heterogeneity in the metabolic function.

To investigate the potentially differential enzymatic activities among individuals affecting their drug-related metabolic capacity, the inference of genetic variants from transcriptomics data was performed using *Long Ranger* (10X Genomics). This set of pipelines processes 10X Chromium sequencing output to align reads and call and phase SNPs, indels and structural variants [412]. To obtain single-cell transcriptomics data, PHHs from two donors (previously used for scATAC-seq analysis) were incubated in the four experimental conditions (DMSO, Cocktail, FFA or FFA+Cocktail) and scRNA-seq performed using the 10X Genomics platform (Methods). Thereafter, the four libraries per donor were prepared and quality control was performed in a Bioanalyzer 2100 (**Figure 39**). This batch was performed with Dr. Ioannis Deligiannis (Helmholtz Pioneer Campus, Helmholtz Zentrum München). *Long Ranger* (10X Genomics) was then used to infer SNPs from each of the eight samples (Methods).

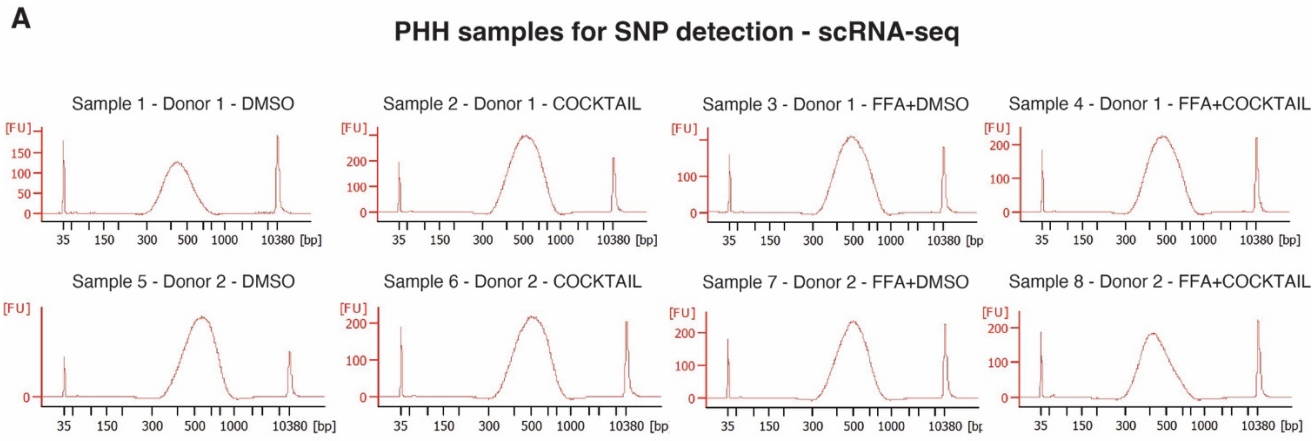


Figure 39. Bioanalyzer electropherograms of the eight PHHs samples, four per donor corresponding to the four treatments, processed for scRNA-seq. *Quality control of the libraries shows the expected traces in bioanalyzer.*

The phased variants outcome file was used to select the chromosome portion where the gene of interest is found, for instance the main five CYP450 enzymes, per each of the four samples belonging to one donor (Methods). Thereafter, the variants detected across all four samples were selected and their biological relevance was verified in the databases: NCBI dbSNP, ClinVar and PharmVar and summarized in **Figure 12**.

A total of eight intron variants were found in donor 1 (Caucasian), in addition to two missense variants, whereas in the case of donor 2 (Hispanic), 5 intron variants were detected. Two missense variants were found in each of the donors, one of them being shared among them (**Table 7**).

Donor 1 (Caucasian) is a carrier of a polymorphism in the *CYP2C9* gene (*CYP2C9*2*, rs1799853), which is located in chromosome 10, position 94942290, where a single base of thymine (T) substitutes a cytosine (C) nucleotide, representing a missense variant, leading to a change of arginine to cysteine amino acids at codon 144 [319, 413, 414]. Lower functional *CYP2C9* enzymatic activity has been reported in the presence of this mutation [415], resulting in reduced drug clearance, showing the importance of the detection of these mutations and their high clinical impact. Interestingly, the allele frequency of this variant in the Hispanic population has been shown to be higher than that found in Caucasian, Asian, African-American, and Jewish- ethnicities [416]. In this dataset, the Caucasian donor showed a lower *CYP2C9* gene expression compared to the Hispanic donor not carrying this missense variant (**Figure 40**). However, a lower overall expression of the five CYP450 enzymes in the Caucasian donor was observed.

Table 7. Summary of present SNPs in the two human donors. (Source: dbSNP database from NCBI, ClinVar and PharmVar).

Donor	Ethnicity	Cytochrome	Chromosome	Position	Variant	Alleles	rsID	Type
1	Caucasian	CYP2C9	**chr10**	94942290	T	C>T	*CYP2C9*2,* rs1799853	Missense variant
1	Caucasian	CYP2C9	**chr10**	94942538	A	G>A	rs2860905	Intron variant
1	Caucasian	CYP2C9	**chr10**	94947445	T	C>T	rs4086116	Intron variant
1	Caucasian	CYP2C9	**chr10**	94950236	C	T>A or T>C or T>G	rs2984310	Intron variant
1	Caucasian	CYP2C9	**chr10**	94952643	G	A>G	rs2475376	Intron variant
1	Caucasian	CYP2C9	**chr10**	94972974	G	T>A or T>G	rs1856908	Intron variant
1	Caucasian	CYP2C9	**chr10**	94982060	G	A>C or A>G	rs1934968	Intron variant
1	Caucasian	CYP2C19	**chr10**	94804000	A	G>A	rs4494250	Intron variant
1	Caucasian	CYP2C19	**chr10**	94821337	G	A>C or A>G	rs10786172	Intron variant
1	Caucasian	CYP2C19	**chr10**	94842866	G	A>C or A>G	*CYP2C19*4,* rs3758581	Missense variant
2	Hispanic	CYP1A2	**chr15**	74755085	C	T>C	rs2470890	Intron mutation
2	Hispanic	CYP2C9	**chr10**	94952643	G	A>G	rs2475376	Intron variant
2	Hispanic	CYP2C9	**chr10**	94981151	C	T>C	rs9332197	Intron variant
2	Hispanic	CYP2C9	**chr10**	94982060	G	A>G or A>G	rs1934968	Intron variant
2	Hispanic	CYP2C19	**chr10**	94821337	G	A>C or A>G	rs10786172	Intron variant
2	Hispanic	CYP2C19	**chr10**	94842866	G	A>C or A>G	*CYP2C19*4,* rs3758581	Missense variant
2	Hispanic	CYP2C19	**chr10**	94849811	C	T>C	rs4917623	Intron variant
2	Hispanic	CYP2D6	**chr22**	42126722	T	G>T	rs79392742	Missense variant

Another example is the mutation found in donor 1, in the *CYP2C19* gene (*CYP2C19*4*, rs3758581, I331V) [319], in chromosome 10, position 94842866 (**Table 7**) alternative allele is a guanine (G) replacing an adenine (A) nucleotide, entailing a missense mutation changing isoleucine by a valine amino acid in codon 331 resulting in loss of function of the protein [413, 414, 417]. Therefore, clinical *CYP2C19* testing should take the potential presence of relevant SNPs into consideration, as the allele contains both gain-of-function [c.−806C>T (**17*)] and loss-of-function [c.1A>G (**4*)] (associated with poor metabolizers) variants on the same haplotype [418].

This *CYP2C19* mutation (*CYP2C19*4*, rs3758581, I331V) was as well present in donor 2 (**Table 7**) The alternative allele frequency among different ethnicities such as African, American, Hispanic, East and South Asian, and European range between 89% (South Asian) and 100% (African) [419]. Therefore, finding a variant in the two donors could be expected.

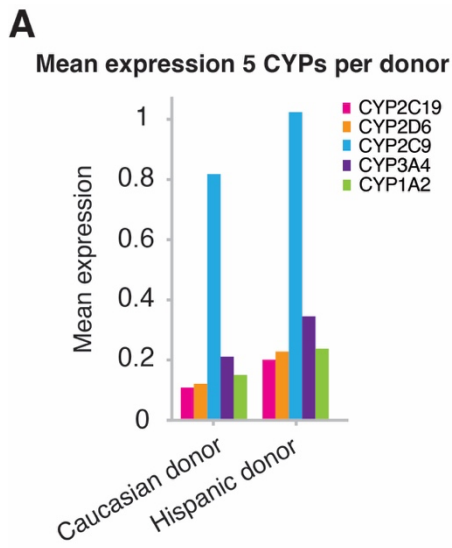


Figure 40. The presence of SNPs might impact the inter-donor CYP2C9 variability. (A) *Bar plot depicting the mean expression of the main five CYP450 enzymes in the Caucasian and the Hispanic donors.*

A second missense mutation in was inferred in donor 2 (Hispanic), in the *CYP2D6* gene (rs79392742, A449D) in chromosome 22. Precisely located in position 42126722, where a thymine (T) base substitutes guanine (G) resulting in the exchange of alanine by aspartic acid amino acids (**Table 7**) [316, 319, 413, 414]. This rare variant has been recently reported to display decreased enzymatic activity (44.4 ± 6.9% vs. 100%) of the CYP2D6 protein due to heme binding perturbation [420]. In another study, this variant was identified in one patient medicated with metoclopramide, resulting in decreased drug clearance and subsequent increased risk for the development of toxic or adverse reactions [421]. This showcases the clinical relevance of the presence of many mutations in drug-metabolizing enzyme genes, not only in *in vivo* studies but also in *in vitro* drug development using PHHs from a single donor.

Taken together, these results showed that the presence of SNPs caused mutations leading to missense variants that affect the enzymatic activity of the genes affected, thus impacting the xenobiotic-related metabolic capacity of the individuals carrying them. Furthermore, these results also show that scRNA-seq performed with the 10X Genomics platform allows for the identification of relevant SNPs, although it being a method that captures the 3´ end of transcripts and not its full length [422]. Therefore,

the analysis of the metabolism of drugs that are metabolized by these CYP isoforms should be taken into consideration in pharmacogenomics and in vitro studies using PHHs isolated from single human donors.

Exploration of factors affecting the drug-related metabolic capacity of primary human hepatocytes *in vitro* at the transcriptomic level

In order to explore other factors that impact the heterogeneous responses of PHHs at the transcriptomic level towards a drug cocktail, chronic fat accumulation or both simultaneously, PHHs from four donors were used for scRNA-seq using a high-throughput droplet-based approach (10X Genomics). Thereby, four experimental conditions optimized in hepatoma cell lines HepG2 and HepaRG were used: Vehicle (DMSO), five-drug cocktail (Cocktail), free-fatty acids (FFA), and the combination of free fatty acids + five-drug cocktail (FFA+Cocktail) (Methods). Two of the donors were previously used in the scATAC-seq (Chapter III) and the SNPs analysis (previous section 4.1). In total, two batches of the same experiment were performed including two donors per batch, and the first batch was performed with Dr. Ioannis Deligiannis (HMGU). The donors were males aged between 18 and 57 years, without liver diseases and a normal Body Mass Index (BMI), non-diabetic, and representing the most common age range commercially available PHHs (Methods).

After a total of 72 h of incubation, 66 of which were exposed (DMSO) or five-drug cocktail, the cells were collected, magnetic cell separation (MACS) for gentle live cell selection was performed, scRNA-seq using 10X Chromium immediately performed (Methods). The application of the 10X Chromium droplet-based approach enabled the analysis of thousands of cells simultaneously, allowing the dissection of cellular characteristics individually, to find subgroups or subtypes of cells in a seemingly homogeneous population. However, the encapsulation of cells larger than 30 µM is challenging due to the risk of clogging the microfluidic channels of the chip [423]. In the second batch, the encapsulation failed for samples 6 and 7, corresponding to the Cocktail-treated cells and the FFA-treated cells from one of the donors, respectively. This limitation is to take into consideration when working with hepatocytes, which range around 30-40 µM in diameter [326, 327], hindering encapsulation effectivity and efficiency.

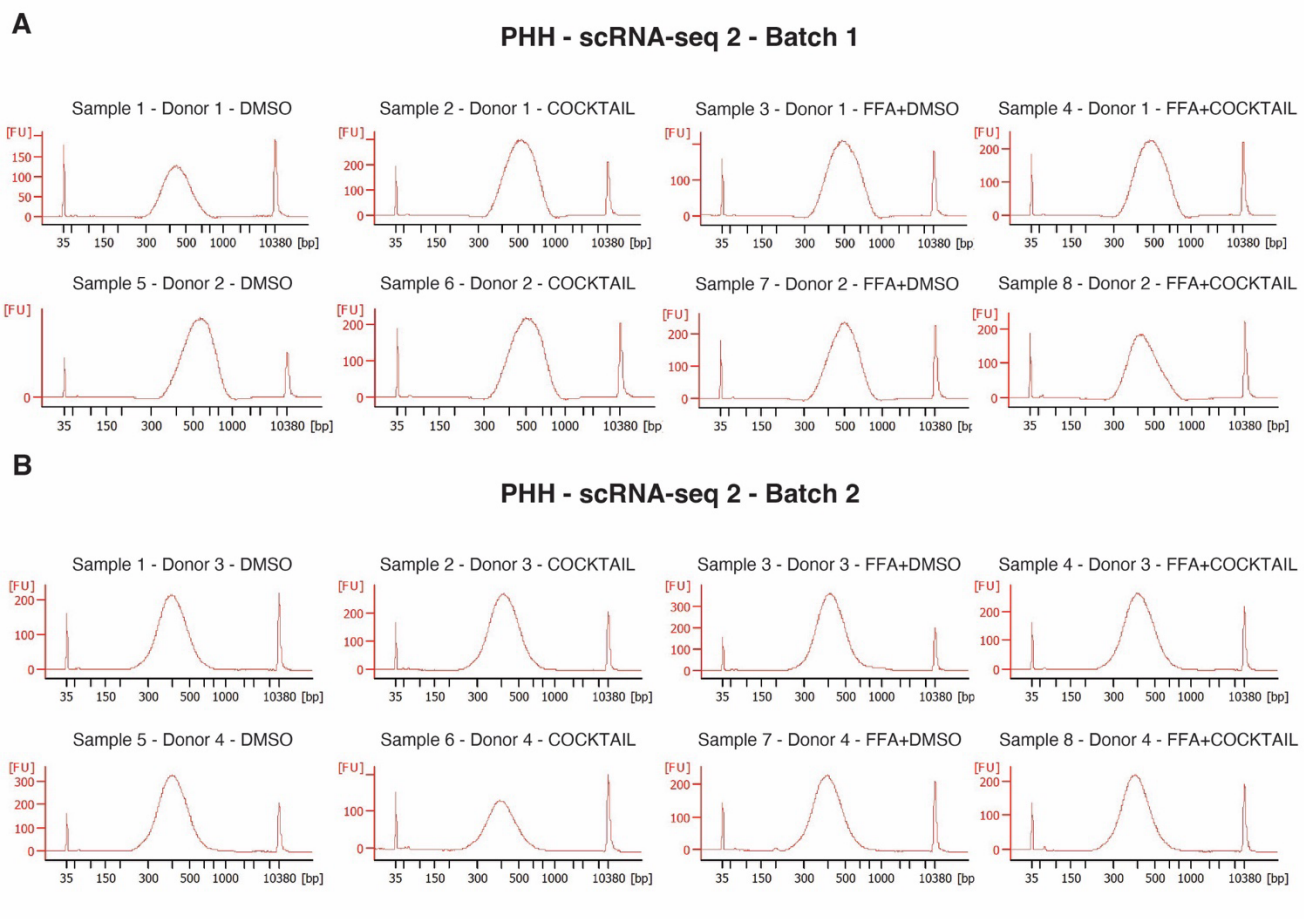


Figure 41. Bioanalyzer electropherograms of the libraries generated in the two experimental batches composing scRNA-seq experiments using PHHs. *Hepatocytes proceeding from each of the donors were incubated in the four treatment conditions in the experiment: DMSO, Cocktail, FFA, FFA+Cocktail.* ***(A)*** *Libraries profiles of the first eight samples comprising batch 1.* ***(B)*** *Libraries profiles of the second eight samples comprising batch 2.*

The electropherogram profiles obtained in the Bioanalyzer 2100 after library preparation and quality control represented the classical Gaussian distribution characteristic of scRNA-seq libraries (**Figure 41**). After assessing the quality of the libraries, these were quantified and sequenced (Methods). The libraries from the first batch were performed with Dr. Ioannis Deligiannis (Helmholtz Pioneer Campus, Helmholtz Zentrum München), and the computational analysis for sections 3.1.1 and 3.1.4 was performed by Ms. Maria Richter (Helmholtz Pioneer Campus, Helmholtz Zentrum München).

Four subgroups of hepatocytes were identified independently of donor and treatment

In vitro cultured PHHs from four donors in four experimental conditions (DMSO, Cocktail, FFA+DMSO or FFA+Cocktail) were assessed for their heterogeneity in the transcriptional responses towards a drug-related challenge such as exposure to a cocktail of five drugs; a metabolic challenge such as intracellular fat accumulation, or both challenges simultaneously by exposing lipid-loaded

cells to a cocktail of five drugs. Four different subgroups of primary hepatocytes were identified, independently of the donor and treatment condition (**Figure 42A**).

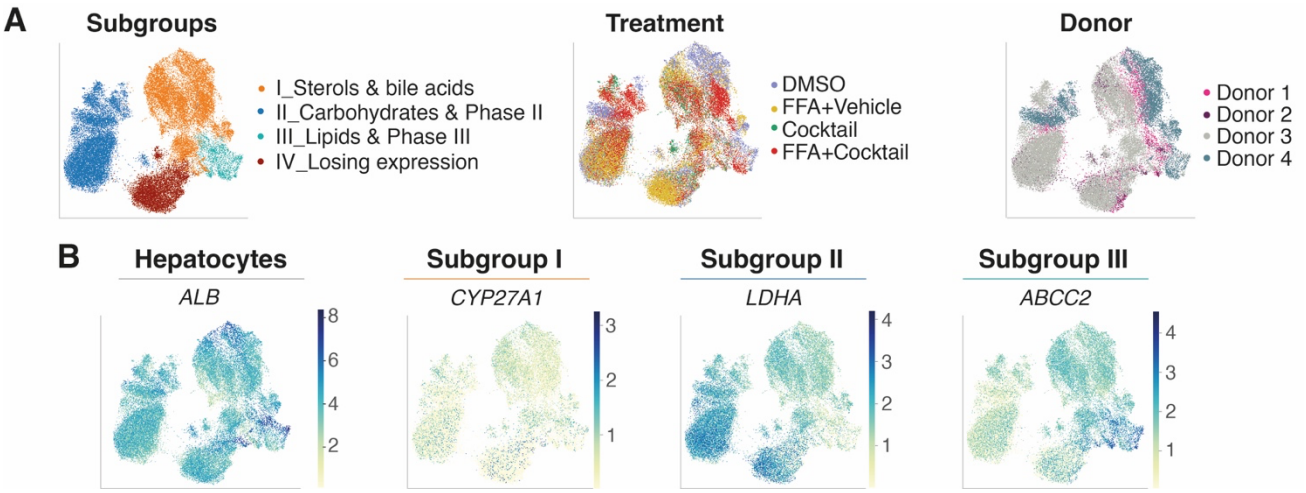


Figure 42. Four distinct subgroups of hepatocytes can be identified independently of donor or treatment. *UMAP colored by* ***(A)*** *Subgroups, treatment, and donor.* ***(B)*** *Marker genes for mature hepatocytes, bile acid metabolism, carbohydrate and phase II metabolism, and lipids and phase III metabolism. Subgroup marker gene expression was grouped by aggregated Louvain clusters. Figure computed by Maria Richter, adapted from Sanchez-Quant & Richter et al. 2023 [1].*

After using *Harmony* [424] for batch correction, the treatment conditions and the donors did not drive the clustering. Markers of differentiated mature hepatocytes (**Figure 42B**), such as albumin (*ALB*), *SERPINA1,* and *TTR* were highly expressed across all cells confirming the enrichment in PHHs in the analyzed dataset.

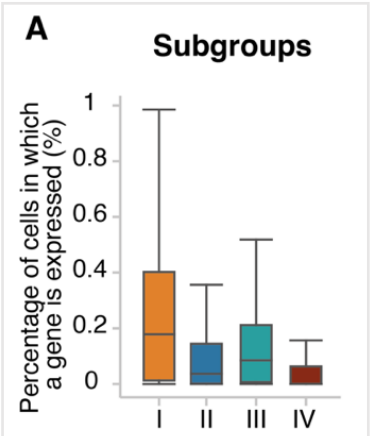


Figure 43. One of the four subgroups was composed of cells characterized by loss in gene expression. ***(A)*** *Box plot showing the percentage of cells expressing a gene per subgroup.*

Among the subgroups, three of them were related to different metabolic pathways in the liver: i) lipids and phase I metabolism; ii) carbohydrates and phase II metabolism and iii) lipids and phase III metabolism (**Figure 42A**). A fourth subgroup was found to be characterized by the downregulation in the expression of key marker genes defining the mature hepatic signature, named losing expression cells (**Figure 42A** and **Figure 43**). It is known that PHHs *in vitro* are phenotypically unstable and undergo a dedifferentiation process after three days in culture [253, 425-431]. Concordantly, in this data, a subgroup of hepatocytes showing loss of hepatic phenotype was detected, compared to the other 3 metabolically active subgroups (**Figure 43**).

Next, to investigate the relevance of these subgroups *in vivo*, a comparison of the gene expression of our three metabolically active subgroups with an *in vivo* dataset generated by Aizarani *et al.* in 2019 [2] was performed (**Figure 44**).

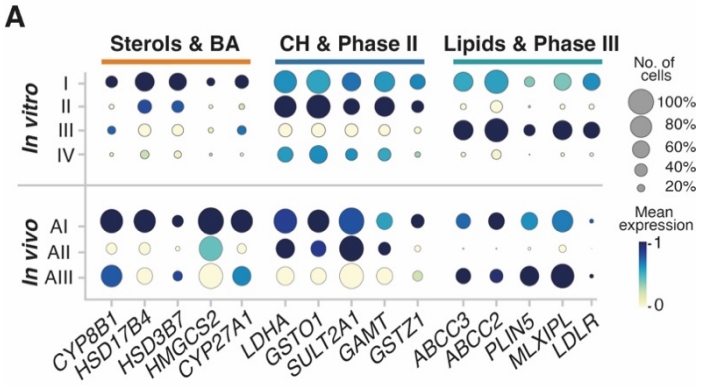


Figure 44. The three metabolically active subgroups identified in vivo could be also observed in the in vivo dataset from Aizarani et al. 2019 [2]. ***(A)*** *Dot plot depicting marker gene expression of hepatocyte subgroups identified in vitro (top) and in vivo (bottom). Dot size: fraction of cells in the group; color scale: mean expression in the group.*

In this study, nine human samples from patients without chronic liver disease (defined as liver damage lasting at least six months) were characterized. When comparing the analysis of these 5 donors with our *in vitro* dataset, we could observe that subgroups I, II, and III could be found *in vivo*, preserving functional specialization in the absence of zonation (**Figure 44**).

Hepatocyte subgroups show diverse transcriptional responses to a phenotyping drug cocktail

To explore the cellular heterogeneity in the drug-related responses among the identified PHHs subgroups, the metabolic profile of the different hepatocyte subgroups was assessed by evaluating their response to the exposure to several drugs. For that, incubation with the phenotyping five-drug cocktail (Sanofi-Aventis cocktail) ("Method optimization" section in Chapter III) [285, 287] was performed for 66 hours (Methods).

Upon Cocktail incubation, the induction of the mRNA of the five CYPs responsible for the metabolism of the five-drug cocktail was monitored showing their upregulation in pseudobulk. A significant upregulation of *CYP2C9* and *CYP3A4* levels upon Cocktail treatment vs. DMSO level was observed (**Figure 45**A). Remarkably, differences in the upregulation levels were observed between hepatocyte subgroups, which were concealed in pseudobulk analyses (**Figure 45**A). For instance, in pseudobulk,

CYP2C9 was upregulated 1.7-fold in Cocktail, while in subgroup III only a 1.1-fold change was detected. Similarly, a 4-fold increase was detected for *CYP3A4* in subgroup I; however, it was not significantly upregulated in subgroup III (**Figure 45**A). A down-regulation of the five targeted cytochrome enzymes was not detected in pseudobulk, nor per subgroup. This indicated that the probe drugs in the cocktail did not have an inhibitory effect on any of the cytochromes at the concentration used in the cocktail. This could have been the case, given that metoprolol has been reported to inhibit *CYP2D6* and *CYP3A4* expression, however without affecting CYP2D6-mediated midazolam metabolism [301]. In addition, omeprazole has been shown to inhibit *CYP2C19* and *CYP3A4* expression in a time-dependent manner [153, 294].

Additionally, differentially expressed genes (DEGs) upregulated specifically per subgroup were also identified. These were not detected as significantly upregulated in the pseudobulk analysis (**Figure 45**B). A similar number of genes were specifically upregulated in every metabolically active hepatocyte subgroup: 122 genes in subgroup I, 102 genes in subgroup II, and 126 genes in subgroup III. This indicates that the three metabolically active subgroups that conserve their mature hepatocyte phenotype possess a distinct transcriptomic signature, with a comparable number of specific differentially expressed genes (DEGs). On the other hand, only 64 genes were specifically upregulated in subgroup IV (**Figure 45**B). This subgroup corresponds to cells identified as losing their characteristic hepatocyte-like gene expression and undergoing a decay in global gene expression along the time in culture (**Figure 43**).

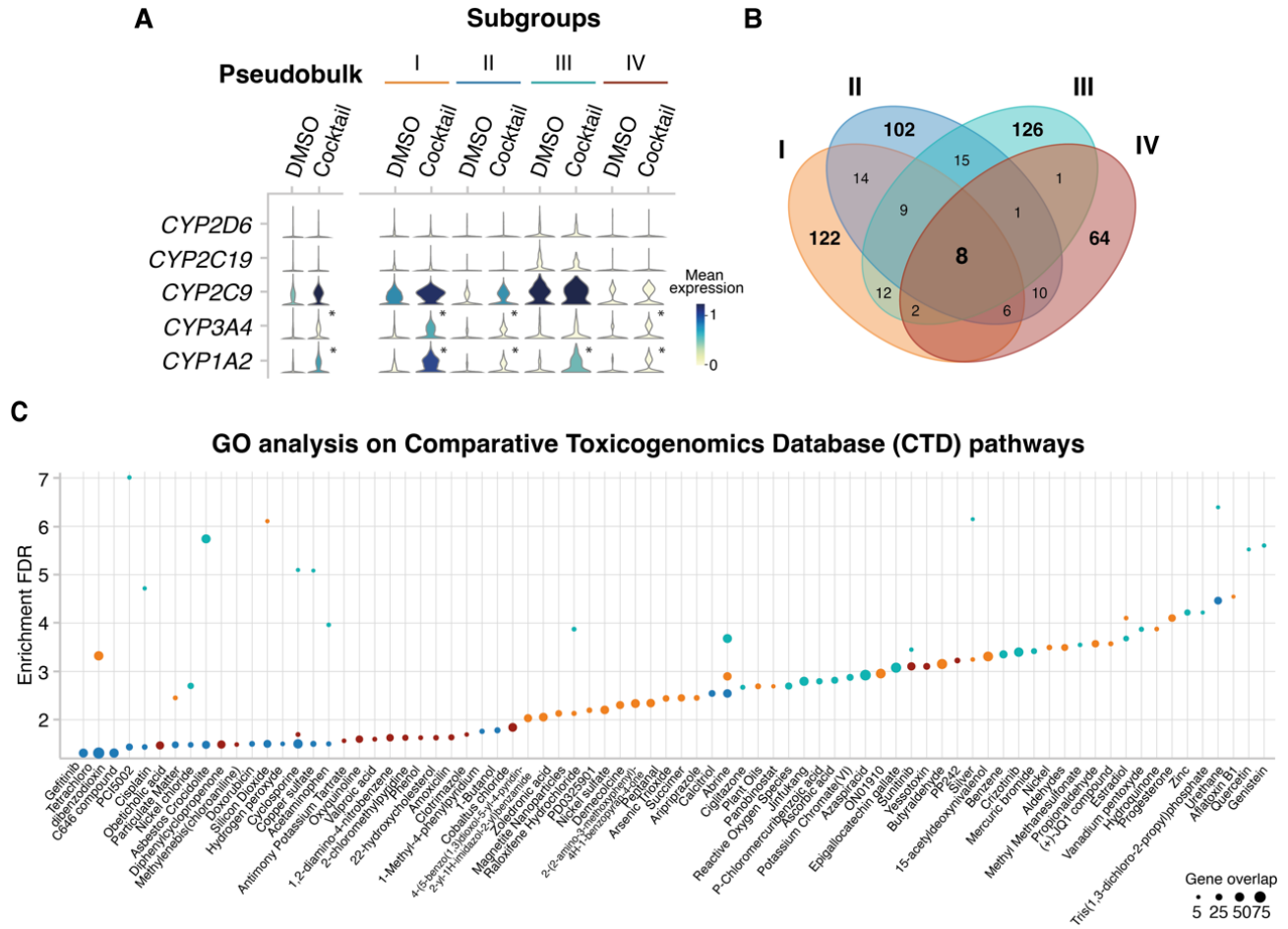


Figure 45. Five-drug cocktail incubation triggers a differential transcriptomic response in the PHH subgroups. ***(A)*** *Violin plot showing the expression levels of the five targeted CYP450 enzymes by the drug cocktail comparing DMSO vs. Cocktail treatment, in pseudobulk and at single-cell resolution per subgroup (*=p-value<0.05 and |log2-fold change|>1, t-test);* ***(B)*** *Venn diagram depicting the overlap of significantly upregulated genes in each subgroup and shared among al subgroups upon Cocktail treatment;* ***(C)*** *Scatter plot showing the enrichment of the genes specifically upregulated per subgroup in pathways known to be involved in the metabolism of the chemical compound specified in the X axis (Drug.CTD database). Dot size represents the number of overlapping genes in a given pathway. Figure computed by Maria Richter, from Sanchez-Quant & Richter et al. 2023 [1].*

In order to explore toxicological interactions specific to each of the subgroups, Gene Ontology (GO) analysis of subgroup-specific DEGs was performed using a database suited for drug-disease or drug-phenotype interactions exploration (Comparative Toxicogenomics Database) [125, 432, 433]. This comparison showed that based on their differential transcriptomic profile, each hepatocyte subgroup is specialized in the metabolism of certain xenobiotics (**Figure 45**C). For instance, some compounds such as the antitumorigenic drug sunitinib [434, 435] were only enriched in two subgroups (I and IV). Only the metabolic pathway for abrine was upregulated in all of the metabolically active subgroups (I, II and III). Finally, the pathways for the metabolism of other compounds were only enriched in one subgroup, such as quercetin and gefitinib were enriched in subgroups I and II, respectively.

In brief, these results indicate that hepatocyte subgroups showed differential transcriptional responses upon a metabolic challenge such as the exposure to a five-drug cocktail. These were characterized by both subgroup-specific transcriptional profiles as well as shared metabolic pathways. Additionally, a different potential for metabolizing chemical substances such as endobiotic (endogenous) and xenobiotic (exogenous) could be identified among the four subgroups.

Intracellular lipid storage alters the transcriptional variability of hepatocyte subgroups

It is known that changes in CYP450 activity correlate with altered lipid metabolism [207, 225, 354]. At the single-cell level, the effect of hepatic steatosis has been found to change the transcriptomic profile of parenchymal and non-parenchymal cells, as well as the cellular composition of the liver [106, 210, 211, 355]. Specifically in hepatocytes, lipid metabolism is disrupted upon fat accumulation due to the alteration of key enzymes in the lipid synthesis, storage, and clearance pathways [106, 356, 357]. Moreover, increased chemokine production has been observed associated with the inflammation process happening in NAFLD [358, 359].

To explore the effect of lipid accumulation on the metabolic capacity of metabolically functional subgroups of PHHs, hepatic steatosis *in vitro* was modeled by incubating the cells with oleic and palmitic acid (200 µM mixture in a 2:1 ratio) [260], recreating benign chronic steatosis with minimized lipotoxicity and apoptotic effects [262, 436] (Methods). The PHHs were loaded with FFA for 72 h and their individual transcriptional profile was analyzed (Methods).

In order to characterize the identified subgroups when exposed to intracellular lipid accumulation, differential gene expression analysis was performed. Among the top five DEGs upon chronic accumulation of lipids per metabolically active subgroups (I, II and III), lipid metabolism and storage-related genes were detected (**Figure 46**A).

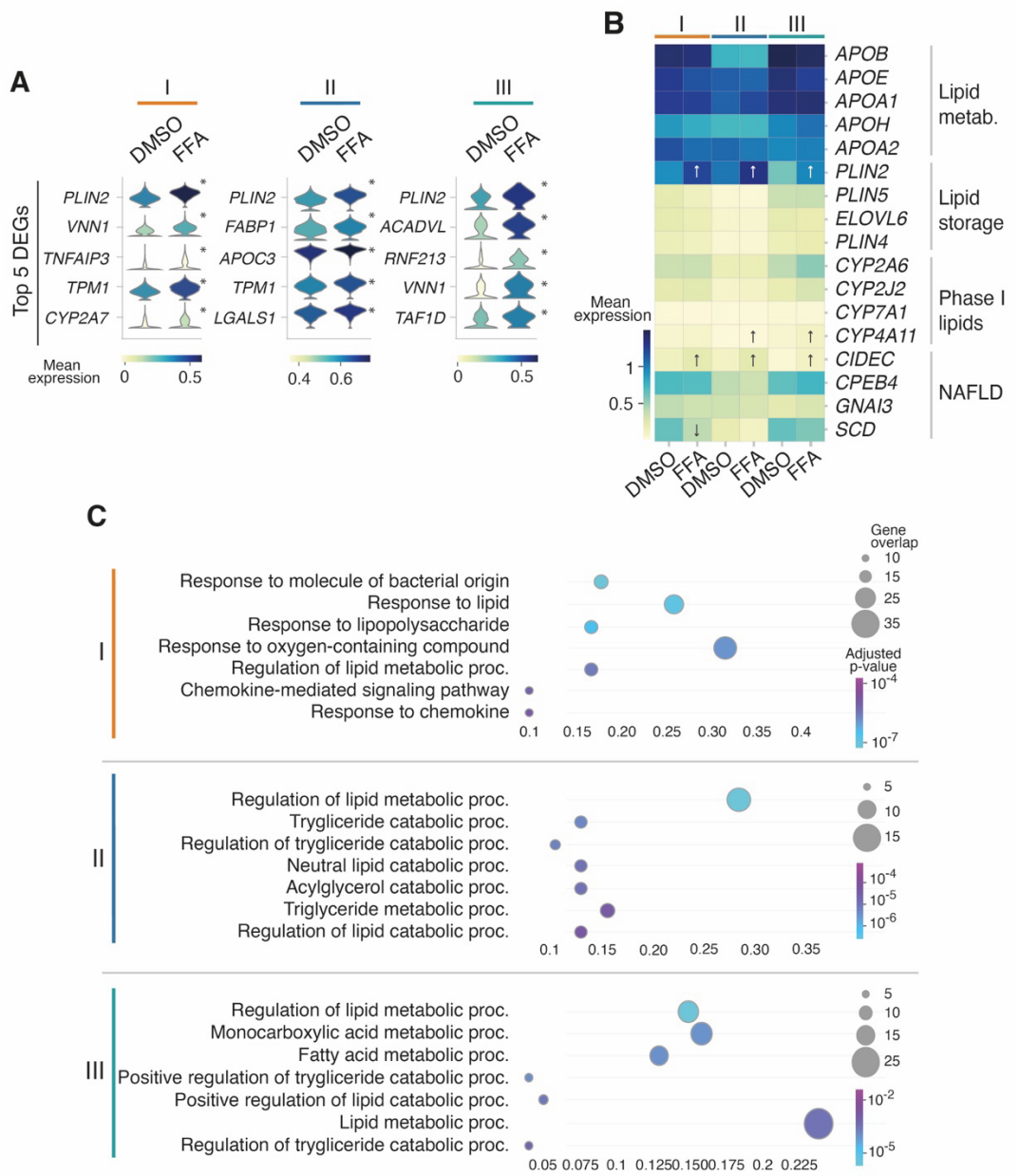


Figure 46. Hepatocyte subgroup-specific metabolic signatures are present in responses to intracellular lipid accumulation. (A) *Stacked violin plot depicting the top 5 DEGs upon fat accumulation in subgroups I, II and III in DMSO and FFA treatments (*=p-value<0.05 and |log2-fold change|>0.75, t-test);* ***(B)*** *Heatmap showing the log mean expression of genes related to lipid metabolism, storage and NAFLD-related in DMSO and FFA-treated cells per subgroup. The arrows indicate either up or downregulation, t-tested). Color scale: Mean expression.;* ***(C)*** *Scatter plot depicting the top 7 GO terms of the genes upregulated in each of the metabolically active subgroups upon fat accumulation. Figure computed by Maria Richter, from Sanchez-Quant & Richter et al. 2023 [1].*

For example, the lipid droplet-associated protein *PLIN2* was significantly upregulated in all three metabolically active subgroups, and it has been shown to be related to diet-induced NAFLD [437-439]. In addition, the inflammation marker *TNFAIP3* has also been reported to ameliorate NAFLD and play a protective role against its progression [440, 441], was found to be upregulated in subgroup I (**Figure 46**B).

Further analysis of the lipid metabolism in the three metabolically active subgroups upon fat accumulation was performed, interrogating the expression of representative genes (**Figure 46**B). For instance, *PLIN2* and *CIDEC*, triglyceride accumulation in lipid droplet-related genes, were significantly upregulated upon fat accumulation in the three metabolically active subgroups. Similarly, *CYP4A11* involved in NAFLD progression by inducing ROS-related lipid peroxidation and inflammation, was upregulated in subgroups I and III upon fat accumulation.

For the identification of the effect of fat accumulation on the main biological processes associated to each metabolically active subgroup, gene ontology (GO) analyses using ShinyGO [442] were performed using the significantly upregulated genes per metabolically active subgroup. Subgroup I showed gene overlaps in pathways related to cellular response to lipids, together with lipopolysaccharide and chemokine metabolism [443] (**Figure 46**C). Subgroup II, exhibited a high overlap of genes involved in the regulation of triglyceride metabolic processes, as well as in the acylglycerol catabolic process, indicating that these cells were involved in the clearance of neutral lipids [444, 445]. PHHs in subgroup III showed enrichment in lipid, monocarboxylic acid, and fatty acid-related metabolic processes and lower transcriptional variability, most likely due to their more coordinated response to fat accumulation.

Next, to investigate how chronic fat accumulation affects the expression of the main cytochromes involved in the xenobiotic metabolism, their expression per subgroup was depicted (**Figure 47**). Upon FFA incubation, the expression of CYP450 cytochrome enzymes involved in drug and endobiotic substance metabolism was affected differently across the four subgroups. For instance, the expression level of CYP2D6 and CYP1A1 was significantly upregulated in subgroup III upon FFA treatment. The expression of CYP2C9 and CYP3A5 was increased in subgroups I and III upon fat accumulation compared to DMSO level and decreased in subgroups II and IV (**Figure 47**). Lastly, the expression of *CYP1A2* was increased in subgroups I, III, and IV, however decreased in subgroup II. This indicates that chronic intracellular lipid accumulation might lead to distinct transcriptional effects in different hepatocyte subpopulations, hence

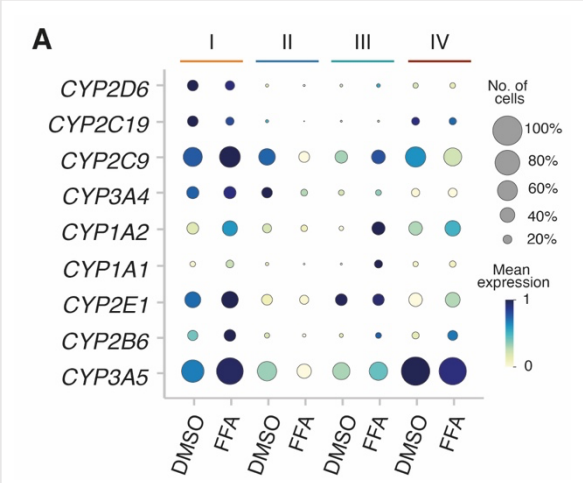


Figure 47. Fat accumulation affects the expression of endo- and xenobiotic metabolizing enzymes. (A) *Dot plot depicting the expression of the main CYP450 enzymes per each of the four subgroups. Dot size: fraction of cells in the group; color scale: mean expression in the group.*

impacting the expression of CYP450 enzymes and thereby the drug-related metabolism capacity in a different manner in each subgroup.

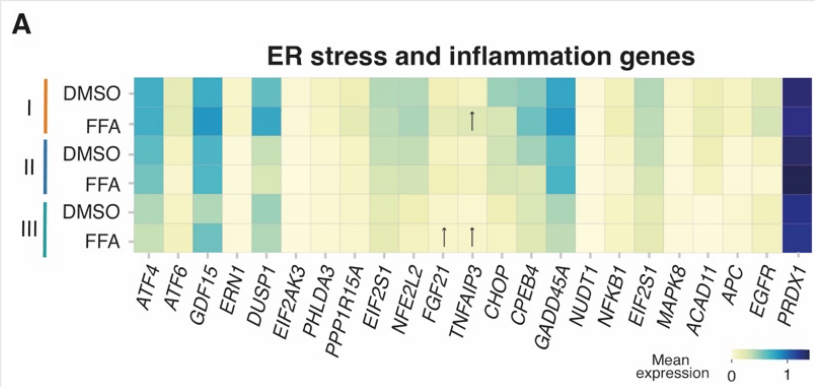


Figure 48. Endoplasmic reticulum-related stress and inflammation pathways were activated by fat accumulation in subgroup I of hepatocytes. *(A) Matrix plot depicting ER stress, oxidative stress, and inflammation-related genes in the three metabolically active subgroups in DMSO vs. FFA treated cells. Color scale: Mean expression.*

Since hepatic lipid accumulation and NAFLD have been associated to oxidative stress, endoplasmic reticulum (ER) stress, and inflammation processes [359, 446-448], the expression of key marker genes in these pathways was investigated (**Figure 48**). The ER stress-responsive transcription factors *ATF4* and *ATF6*, as well as *GDF15* and *PPP1R15A* showed a higher mean expression in subgroup I compared to the other two subgroups upon intracellular fat accumulation. The master regulator of oxidative stress *NFE2L2* was also showing a higher mean expression in subgroup I upon fat accumulation (**Figure 48**). The peptide hormone *FGF21* showed a significant upregulation in Subgroup III upon FFA treatment compared to DMSO treatment. This hormone has been reported to increase its level upon hepatic lipid accumulation, obesity, and NASH; and to be protective against hepatic lipotoxicity [449]. In addition, *TNFAIP3* which has been reported to have a protective role against NAFLD and its progression [440, 441]was significantly upregulated upon fat accumulation in subgroups I and III, corresponding to sterols and phase I metabolism, and lipids and phase III metabolism, respectively.

This indicated that fat accumulation affected the phase I metabolic capacity of primary hepatocyte subgroups differently. Moreover, it increased the ER- and oxidative stress-related response in a different manner between subgroups, affecting mostly subgroup I. Specifically in subgroup III, this response was attenuated, and protective genes against lipid accumulation were significantly upregulated upon FFA treatment.

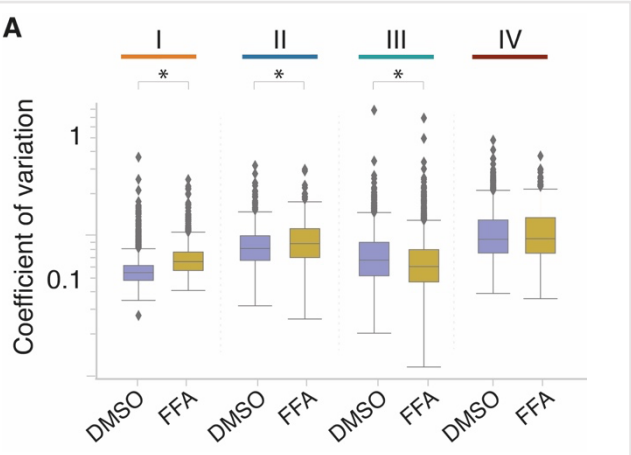


Figure 49. Intracellular fat accumulation significantly increases transcriptional variability in subgroups I and II and decreases it in subgroup III. ***(A)*** *Box plot showing the coefficient of variation per gene calculated per subgroup upon DMSO or FFA treatment (*=p-value<0.05, MannWhitneyU). Figure computed by Maria Richter, from Sanchez-Quant & Richter et al. 2023 [1].*

Furthermore, the coefficient of variation for DMSO- and FFA-treated cells was calculated to measure transcriptional variability. Thereby the coordinated transcriptional response triggered by fat accumulation on PHHs in each subgroup can be assessed. The cells losing expression (subgroup IV) showed the highest transcriptional variability (**Figure 49**). In functional subgroups I and II, lipid accumulation significantly increased the variability, but it was found to be significantly diminished in subgroup III (**Figure 49**). This indicates that the latter responded more coordinately towards the accumulation of lipids [16, 22, 185, 187, 450].

Taken together, chronic intracellular lipid accumulation increased transcriptional variability in metabolically active hepatocyte subgroups I and II, potentially affecting the fine-tuned regulation of lipid metabolism. Moreover, ER stress-related genes were more highly expressed in subgroup I, responsible for the metabolism of sterols and phase I, upon fat accumulation. Importantly, in subgroup III, specialized in the metabolism of lipids, transcriptional variability was reduced upon fat accumulation suggesting a robust and tight coordinated response to chronic accumulation of lipids, together with a significant upregulation of NAFLD-ameliorating genes.

Aging increases cell-to-cell variability among PHHs and induces stress and inflammation

Aging has been shown to increase the transcriptional variability among a seemingly homogeneous population of cells, and this increase in heterogeneity prompts the discoordination that occurs in the aging phenotype across tissues and organisms [16, 185-187, 406, 451]. Using the approximately 30-year age difference between young and old individuals, the effect of aging at the transcriptomic level in single PHHs was investigated. The donors in this study belonged to two different age groups: two of them were young at 18 and 26 years of age, whereas the other two were older, being 56 and 57 years of age (Methods). In total, 11,279 young and 26,953 old cells were studied, and their transcriptome was compared to perform preliminary analyses.

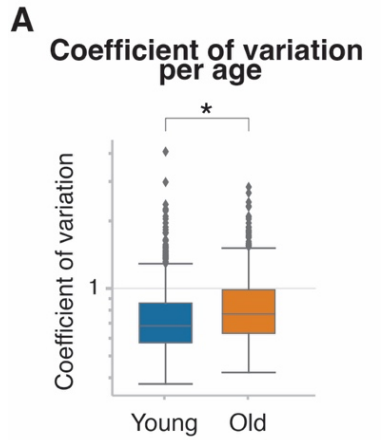


Figure 50. Aged cells show increased transcriptional variability compared to young cells. ***(A)*** *Box plot showing the calculated coefficient of variation for young vs. old cells with an age difference of 30 years. (*=p-value<0.05, MannWhitneyU).*

To explore the effect of aging on the coordinated transcriptional response of PHHs and on cell-to-cell variability, the coefficient of variation of young vs. old cells was calculated (Methods, **Figure 50**). A 1.13-fold increase in the value was observed in older cells, suggesting that aging could increase the variability intrinsic in a population of PHHs. In concordance with a body of literature showing that increases the transcriptional noise and cell-to-cell variability, leading to less coordinated transcriptional responses [111, 184-186, 452, 453].

Additionally, a decrease in the number of genes detected per cell has been recently reported in mouse liver and in other tissues associated to aging [188]. Therefore, the impact of the donor age on the number of genes detected in old vs. young cells, as well as the number of counts per cell per age was investigated (**Figure 51**). The mean number of genes was decreased from 2,998 to 2,363 in young vs. old cells, respectively (**Figure 51**A). The mean number of counts per age was also reduced from 16,918 to 10,678 in young vs. old cells (**Figure 51**B, C). This indicates that the number of transcripts expressed is decreased in aging hepatocytes, in concordance with previous literature [188].

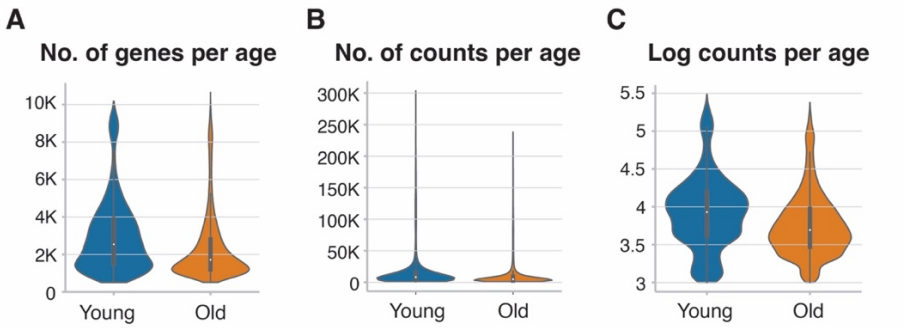


Figure 51. A decreased number of genes and counts was detected in old vs. young cells. ***(A)*** *Number of counts per cell;* ***(B)*** *Number of genes per cell;* ***(C)*** *Number of counts per cell, per young and old donors.*

Batch correction for the integration of the two batches with two donors each was performed using *Harmony* [424]. Cell clustering was not driven by donor age nor by treatment (**Figure 52**). The two

batches performed merged into the UMAP space, and the mitochondrial fraction was low across all cells in the study.

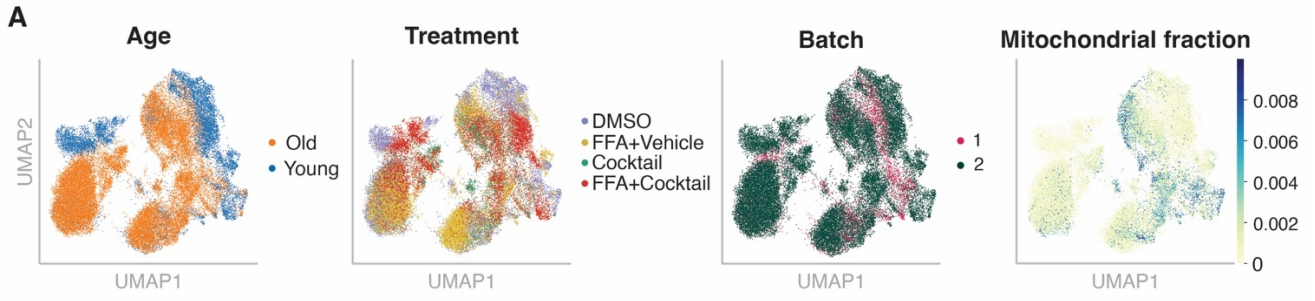


Figure 52. Age, treatment condition or batch did not drive cell clustering. (A) *UMAPs colored by clustering of young and old cells; treatment, batches processed in the experiments and the mitochondrial fraction across cells.*

Next, the contribution to each of the four subgroups by young or old cells was explored (**Figure 53**). In each of the three metabolically active subgroups (I, II and III), the percentage of young cells was lower than the old cells. Young cells represented 60.3% in subgroup I; 43.3% in subgroup II and 70.9% in subgroup III. The cells losing expression (subgroup IV) showed less than 30% of young cells (**Figure 53**). The presence of a lower number of young cells in the dataset can be explained as a technical artifact by the failure in encapsulation in the Chromium chip, leading to samples 6 and 7, corresponding to two of the young samples being barcoded together instead of in single cells. The highest percentage of old cells was captured in subgroup IV, which corresponded to losing expression cells (**Figure 53**).

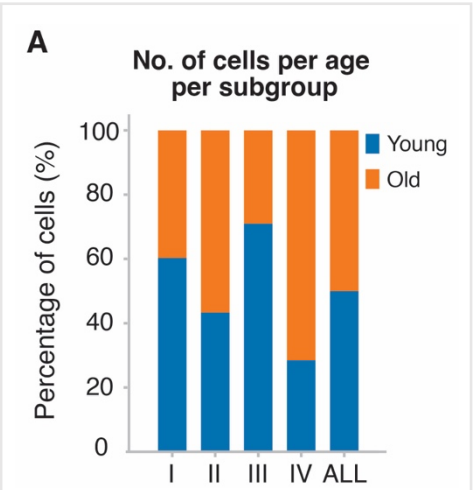


Figure 53. The age composition of the four subgroups shows the lowest number of young cells in the subgroup of cells losing expression. (A) *Bar plot depicting the percentage of young and old cells per subgroup.*

It is known that aging works to the detriment of the metabolic capacity of the liver. A decrease in the blood volume and flow in the organ and in the lipid and xenobiotic metabolism has been described [305]. Aging affects hepatocyte function [454] and CYP enzyme activity [208, 408, 455-457]. The metabolism of drugs in the study was assessed using the

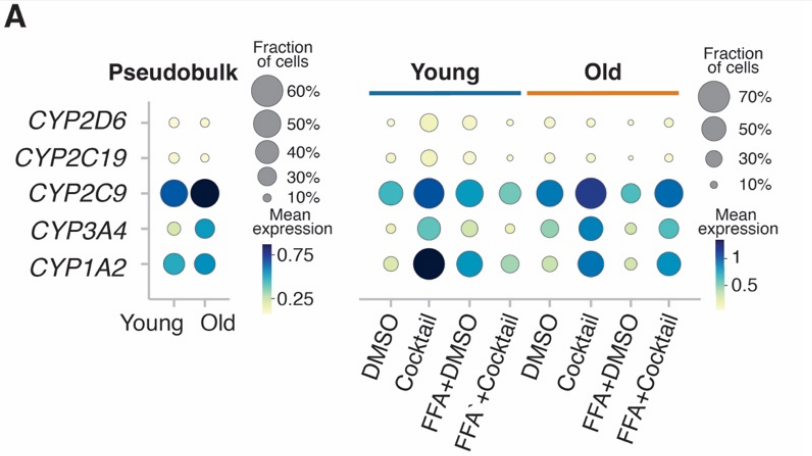


Figure 54. Aging affects the drug-related metabolic capacity of hepatocytes in pseudobulk and per subgroup. (A) *Dot plot showing the expression of the main five CYP450 isoforms in young vs. old cells in the four treatment conditions. The dot size represents the percentage of cells expressing that gene, and the color scale the mean expression level.*

well-characterized cocktail approach (Sanofi-Aventis) and then comparing young vs. old cells (Methods, **Figure 54**). Therefore, the expression of the inducible CYP450 isoforms that metabolize the five probes composing the drug cocktail, in pseudobulk analysis and at single-cell resolution was investigated. In pseudobulk, a significant upregulation of *CYP2C9* and *CYP3A4* expression levels was detected (**Figure 54**). However, at single-cell resolution, subtle changes in the proportion of cells expressing the marker genes and the mean expression level could be identified, which were otherwise concealed in the averaged measurements displayed in pseudobulk analyses (**Figure 54**). In FFA+Cocktail treatment, the expression of the five CYP450 enzymes targeted by the drug cocktail is higher in old hepatocytes compared to young, particularly for *CYP2C9, CYP1A2* and *CYP3A4,* where a 1.74-fold, a 1.88-fold, and a 3.74-fold increase were detected, respectively (**Figure 54**). When comparing the expression levels upon Cocktail treatment vs. FFA+Cocktail treatment within cells of the same age, fat accumulation in older cells caused a milder downregulation of the CYP expression levels, compared to that within young cells. While the sample size in terms of the number of donors, but not in the number of cells per age, prompts caution when interpreting these results, the data suggests that aging induces a less coordinated transcriptomic response in aged PHHs vs. young cells. Furthermore, changes in gene expression can be observed when comparing young vs. old cells at the transcriptomic level upon five-drug cocktail treatment.

Additionally, the activation of inflammatory and stress responses has been described to be a mark in the aging phenotype as a natural response related to the function decay of the different organs and tissues [458-460]. Furthermore, hepatic steatosis has been shown to be a phenotypic trait of the healthy aging liver [173], due to the lipid metabolism impairment and the phenotype of an aged liver resembles that of a liver suffering NAFLD in mammals [356, 461]. This indicates that the fat accumulation occurring in aged livers promotes the inflammatory and stress pathways upregulation.

At single-cell resolution, the expression of stress markers such as *GSTA1,* or *LGALS1* was increased in old vs. young cells (**Figure 55**), whereas the expression of the marker *GCLM,* which is a subunit of the GSH synthase responsible for glutathione synthesis, protecting against oxidative stress [462] showed a decrease in old vs. young cells. Similarly, the expression of the known senescence and oxidative stress regulator *TP53* [463, 464] was downregulated in old vs. young hepatocytes.
Furthermore, in old hepatocytes, the expression of the anti-inflammatory marker *TNFAIP3* or the pro-apoptotic marker *TRAF2* was slightly upregulated (**Figure 55**). Interestingly, *TNFAIP3* which had been described to ameliorate NAFLD and to be protective against disease progression [440, 441] was found among the five top-upregulated genes in subgroup I of PHHs upon FFA treatment when age separation was not performed for analysis (**Figure 46**). Moreover, the lipopolysaccharide binding protein (*LBP*) was significantly upregulated in young vs. aged cells. It has recently been reported that a significant downregulation of liver LBP levels promotes oxidative stress and inflammation, leading to the aggravation of NASH progression [465].

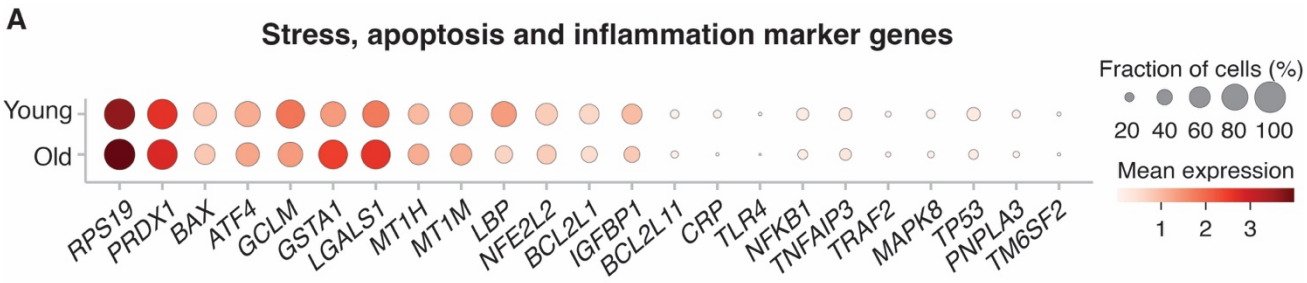


Figure 55. Aging affects the stress, apoptosis, and inflammation transcriptomic response in young vs. old cells. ***(A)*** *Dot plot displaying the expression of stress, inflammation, and apoptosis-related marker genes in young vs. old cells. Dot size corresponds to the percentage of cells. The color scheme corresponds to the mean expression level.*

Altogether, these results indicate that aging impacts the transcriptomic response of PHHs *in vitro* and alters the cellular drug-related metabolic capacity. Stress, apoptosis, and inflammation markers were activated in aged compared to young cells at single-cell resolution. Furthermore, aging increased the transcriptional noise compared to young cells, leading to less coordinated transcriptional responses in PHHs.

Fat accumulation diminishes the drug-related metabolic capacity of hepatocyte subgroups

The simultaneous administration of five or more drugs, known as polypharmacy, is highly common in clinical practice [168, 466]. The intake of several drugs increases the risk of developing hepatotoxicity and

Adverse Drug Reactions (ADRs) such as Drug-Induced Liver Injury (DILI) [467-469]. A factor increasing the risk of DILI is NAFLD, and a higher incidence of DILI has been found in patients suffering from this disease [198, 209 0]. Moreover, at single-cell level, specific transcriptomic response dysregulations have been recently described in NAFLD [470].

Therefore, to assess the impact of fat accumulation on the phase I drug metabolism of PHHs, the previously characterized phenotyping cocktail (Sanofi-Aventis) was used [285]. Changes in the expression of the five CYPs targeted by the drug cocktail were analyzed by comparing Cocktail- vs. FFA+Cocktail-treated cells (**Figure 56**). Fat accumulation decreased the expression levels of the five targeted cytochromes in all subgroups (**Figure 56**). For instance, in subgroups I, II, and IV, *CYP3A4* was significantly upregulated upon cocktail treatment, but its induction was diminished upon chronic lipid exposure. The largest dynamic range of CYP expression was observed for subgroup III. Upon FFA+Cocktail treatment, the expression levels of all five CYPs except for *CYP1A2* were significantly downregulated in comparison to baseline DMSO (**Figure 56**).

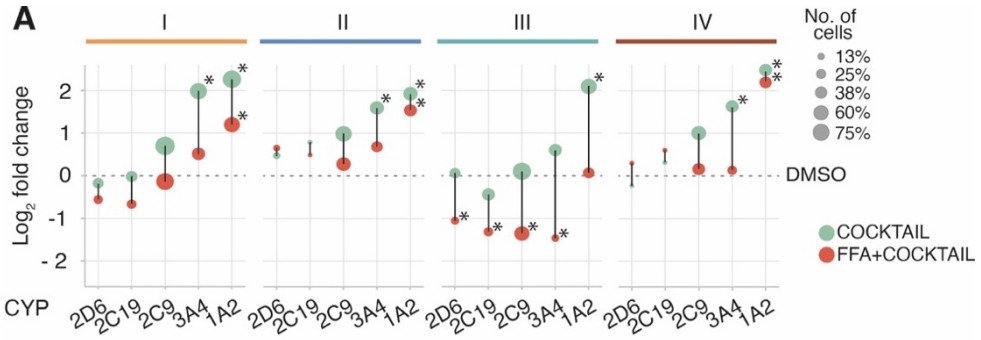


Figure 56. The concomitant treatment of FFA and the five-drug cocktail impairs the drug-related metabolic capacity of the four hepatocyte subgroups. (A) *Scatter plot depicting the log_2 fold-change induction of the five CYP enzymes targeted by the drug cocktail setting DMSO as the baseline level. (*=p-value<0.05 and $|log_2$-fold change$|>1$, t-test). Dot size indicates the number of cells in percentage in which the gene is expressed. Dot color indicates the treatment: green for Cocktail and red for FFA+Cocktail. Figure computed by Maria Richter, from Sanchez-Quant & Richter et al. 2023[1].*

Pursuing a deeper analysis of the transcriptional changes between Cocktail and FFA+Cocktail conditions, differential expression analysis was performed in pseudobulk (Methods). After setting DMSO as a baseline level of expression, 264 genes were upregulated exclusively under Cocktail treatment; 234 genes were commonly upregulated in both Cocktail treatment and FFA+Cocktail; and 602 genes were upregulated specifically in FFA+Cocktail, suggesting the association of a more complex network of biological processes to this condition, considering all cells and per subgroup (**Figure 57**A). The percentage of DEGs belonging to each of the three categories was similar in the four subgroups of hepatocytes, indicating that drug metabolism was affected in a similar manner by

intracellular fat accumulation (**Figure 57**B). To further investigate the affected pathways, GO analyses were performed on these genes (**Figure 57**C).

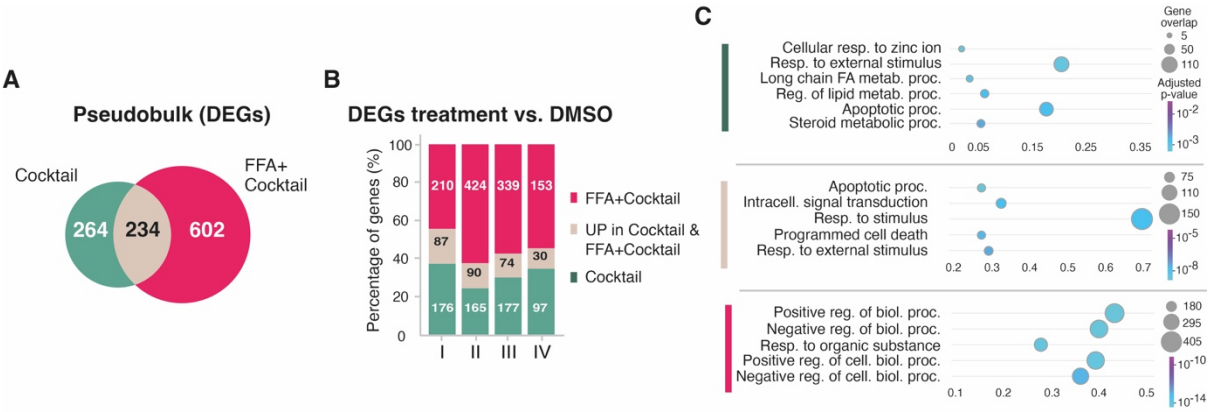


Figure 57. Treatment with FFA+Cocktail activates a complex network of biological processes and stress-related pathways. ***(A)*** *Venn diagram depicting the DEGs in Cocktail (green) vs. FFA+Cocktail-treated cells (red);* ***(B)*** *Bar plot showing for each subgroup the percentages of genes specific to a) Cocktail vs. DMSO (green); b) Genes up-regulated in both, Cocktail vs. DMSO and FFA+Cocktail vs. DMSO (beige); c) Specific to FFA+Cocktail vs. DMSO (magenta);* ***(C)*** *Scatter plot depicting the top five GO terms of the upregulated genes specifically per treatment. In all figures, upregulated genes in Cocktail are represented in green, shared genes in Cocktail and FFA+Cocktail are represented in beige and upregulated genes only in FFA+Cocktail-treated cells in magenta. Figure computed by Maria Richter, from Sanchez-Quant & Richter et al. 2023 [1].*

For Cocktail-treated cells specifically, an enrichment in pathways responsible for the metabolism of xenobiotic compounds was observed (**Figure 57**C, top, green). Genes commonly upregulated upon Cocktail and FFA+Cocktail showed less specificity towards drug metabolism and showed enrichment in general stimulus-response pathways (**Figure 57**C, middle, beige). Finally, genes specifically upregulated in FFA+Cocktail were enriched in stress-related pathways (**Figure 57**C, bottom, magenta). However, in all subgroups the percentage of genes in each category was similar, suggesting that drug metabolism was comparably affected by lipid accumulation in the four hepatocyte subgroups. In summary, intracellular lipid accumulation in each subgroup led to a characteristic impairment of drug-related metabolism, with several metabolic pathways being simultaneously involved and affected. Therefore, the drug-related metabolic specificity is diminished in the presence of chronic fat accumulation in primary human hepatocytes *in vitro*.

Biological relevance of the findings in PHHs at the epigenetic level and at the transcriptomic level

In this last part, a comparison between the findings through scATAC-seq at the epigenetic level on primary human hepatocyte cells *in vitro* and the differential metabolic profiles characterized at the

transcriptomic level (scRNA-seq) is performed. Indeed, some of the characteristic and distinct changes observed in chromatin accessibility among the four experimental conditions could be detected at the transcriptomic level.

Interestingly, *ALB* was found to be differentially accessible upon DMSO treatment, whereas, in the transcriptomics dataset, its expression was similar across all cells and treatment conditions (**Figure 38**, **Figure 42**). Upon Cocktail treatment, all the differentially accessible genes depicted in **Figure 38** were identified as DEGs in the scRNA-seq dataset, including several CYP450 enzymes and the liver-enriched transcription factor *HNF4A [1, 471].* The specific chromatin openness configuration changes observed in PHHs treated *in vitro* with a five-drug cocktail, where regions upstream CYP genes were accessible, modulate the downstream gene expression response into a characteristic metabolic profile, resulting in the upregulation of CYP expression.

Free fatty acids treatment of PHHs causing intracellular fat accumulation triggered a distinct chromatin accessibility conformation, involving genes related to lipid metabolism and accumulation in hepatocytes *in vitro*. Regions upstream of specific lipid metabolism-related genes were differentially open upon FFA treatment. For instance, peaks upstream of the liver-enriched transcription factor *HNF1A* [472], *ABCC3*, and *APOH* were identified as differentially accessible and as DEGs in FFA-treated cells vs. DMSO-treated cells in the scRNA-seq data [1]. Therefore, this distinct epigenetic profile upon lipid accumulation can be also identified downstream, at the transcriptomic level, where a specific response by hepatocytes toward fat accumulation has also been identified.

Lastly, when cells were concomitantly treated with free-fatty acids and a five-drug cocktail (FFA+Cocktail), regions upstream of *CYP450* enzyme genes related to the metabolism of sterols and inflammatory processes were found differentially accessible. Moreover, regions in proximity to solute carriers such as *SLC19A1* or *SLC7A2,* as well as the transcription factor *ONECUT2* were identified as differentially accessible in FFA+Cocktail-treated cells. At the transcriptomic level, these three genes were identified among the DEGs upon FFA+Cocktail vs. DMSO treatment in the scRNA data[1, 375]. This indicates that, upon exposure to both intracellular fat accumulation and a five-drug cocktail, the chromatin accessibility arrangement of PHHs changes to open regions in proximity to genes related to inflammation and cellular transmembrane transport. In addition, the detected changes at the epigenetic level can also be found at the transcriptomic level upon intracellular fat accumulation and incubation with the five-drug cocktail.

In summary, these results indicate a differential and specific chromatin accessibility landscape and configuration along the genome of PHHs upon a metabolic challenge such as the metabolism of a five-drug cocktail, intracellular lipid accumulation, or both simultaneously. Furthermore, these results

reflect the findings from the transcriptomics dataset analyzed in this thesis, where exposure to the different treatment conditions triggered a differential transcriptional metabolic profile in hepatocytes. The methodology here optimized for the nuclei isolation, lysis and tagmentation for scATAC-seq processing using a high-throughput approach allows for the investigation of open chromatin regions regulating the downstream gene expression.

4. Discussion

In the present doctoral thesis: i) The single-nucleus RNA-seq2 (snRNA-seq2) methodology led to the detection of heterogeneity within oxytocin-positive neurons in the murine hypothalamus that are characterized by the distinct gene expression of key neuronal markers. ii) Assessment of the chromatin accessibility configuration (scATAC-seq) revealed changes in the gene expression regulation in primary human hepatocytes associated with treatment with a drug cocktail, fat accumulation, or both simultaneously. This methodology can be used to assess other characteristics of the liver such as polyploidization. iii) Single-cell RNA-seq on PHHs *in vitro* unraveled the existence of hepatocyte subgroups showing heterogeneous transcriptional responses towards a drug cocktail, intracellular fat accumulation, or both concomitantly. In addition, a preliminary assessment of the effect of aging on primary human hepatocytes showed increased transcriptional variability in older cells.

4.1 Single-cell genomics using plate vs. droplet-based approaches and their application to complex tissues like brain

In this chapter of the dissertation, the aim was exploration of the heterogeneity of hypothalamic oxytocin-positive neurons and their role in feeding homeostasis by applying the recently developed snRNA-seq2 methodology. This method relies on the SMART-Seq2 chemistry from a commercially available kit (SMART-Seq v4 Ultra Low input RNA, Takara), with an additional lysis buffer resulting in a two-step lysis of the nuclear membrane [19]. The single-nucleus RNA-seq methodology has been applied in complex tissues such as brain [25-27, 64, 473-476], lung [477], kidney [478-481], heart [482, 483], pancreas [484] and liver [485]. A major advantage of single-nucleus sequencing is that it enables the exploration of samples where obtaining a single-cell suspension is not readily feasible, or from frozen samples. Furthermore, it minimizes the potential gene expression alteration occurring during single-cell dissociation [486, 487]. In general, plate-based methods offer higher flexibility regarding the used lysis buffers, full-length cDNA sequencing, and the addition of control RNA spike-ins [External RNA Controls Consortium (ERCC)] [325, 488, 489]. Furthermore, plate-based methods provide more control and flexibility on the number of cells from different samples enabling more efficient experimental design [28]. As occurring in the experimental design in this thesis, hypothalamic samples obtained from mice fed chow or an obesogenic high-fat-high-sugar diet were sorted in the same 384-well plate, reducing batch effects by processing the different conditions in the same

experiment. Moreover, the inclusion of a quality control well containing a higher number of nuclei than one is possible when using plate-based approaches [490].

Another key difference between plate-based and droplet-based approaches is the manner the mRNA molecules are sequenced [422, 490, 491]. Full-length transcript sequencing strategies applied in plate-based methodologies such as SMART-Seq2 and snRNA-seq2 rely on the use of a template-switching mechanism for reverse transcription [492] and a 25-nucleotide universal 5′ anchor to capture both ends of mRNA molecules [31]. Therefore, full-length transcript sequencing approaches such as SMART-seq2 [31], MATQ-seq [493] or SUPeR-seq provide the ability to interrogate alternative splicing isoforms, gene fusions and mutations. For instance, lowly-expressed transcripts, splicing variants detection, and higher sensitivity have been proven to be preferentially detected by full-length transcript approaches [28, 422]. In essence, both low-throughput, plate-based methods and high-throughput, droplet-based methods mold to address different research questions and experimental needs [494, 495]. The plate-based snRNA-seq2 method is the ideal alternative for the study of rare specific low-yielding tissues or cell types because of the overall lower throughput and higher sensitivity plate-based methods offer [325, 489] with the added value of a more efficient nuclei lysis. Precisely, the number of yielded cells per hypothalamus rounds 600 to 800 neurons, making the application of droplet-based methods not feasible and low-throughput methodologies more suitable. If the starting number of available cells is limited, the minimum number of cells that are required and the capture efficiency of the method are key factors to consider [28]. While droplet-based methods require thousands of cells as input and are suitable for higher-throughput studies, plate-based methods are of advantage when only a few cells are available [28]. As an example, among the high throughput methods, the most used ones are Drop-seq, inDrops, and 10x Genomics. These three methodologies show a recovery rate 2–4%, 75% and 50% of the loaded cells, respectively and they require >200,000, 2,000–10,000 and >1,000 cells as input, respectively [28, 34, 496-498].

The improved snRNA-seq2 methodology has been developed specifically for frozen mouse liver, albeit the results in this dissertation demonstrate its applicability to hypothalamic GFP-labeled oxytocin-expressing neurons from the brain of transgenic (INTACT) mice. It also allowed the study of flash-frozen archived samples that were collected at different time points and stored for later simultaneous processing. This often occurs when sample collection of several mice, and organs is performed simultaneously but cannot immediately be processed in parallel due to complicated logistics or manpower. An example of this scenario is the processing of samples from transgenic mouse lines that undergo FACS purification.

However, the physical and physiological differences between brain and liver tissues require protocol adaptations [499]. A crucial step is overcoming the technical barrier of efficient a) nuclei isolation and b) nuclei lysis. Nuclei isolation was achieved following the procedure published by Krishnaswami *et al.* 2016 [500], which was optimized to isolate nuclei from human brain neurons for FACS sorting. The number of strokes with the tight and the loose pestle and the buffer composition is crucial for the obtention of a single nuclei suspension where the cell membrane is disrupted but the nuclei membrane integrity is not affected. Thereafter, the nuclei are sorted into the microtiter plate containing lysis buffer, and the nuclei are lysed. For nuclei lysis, the snRNA-seq2 methodology entailed a key improvement in the pipeline: the inclusion of a two-step nuclear lysis, resulting in the improved yield of the number of transcripts detected per nuclei [19]. For transgenic murine hypothalamic samples, the snRNA-seq2 methodology was followed for nuclei lysis, leading to the generation of a high-quality dataset that enables the characterization at the transcriptomic level of oxytocin-positive neurons hypothalamic mouse neurons in a complex experimental design.

The absolute number of transcripts has been shown to be proportional to the cellular size, with the mRNA concentration remaining relatively equal among cells despite the considerable variation in absolute mRNA numbers [330]. Cellular size modulates the rate of intracellular biochemical reactions [501, 502] and it has been reported that at single cell level, the number of transcripts has been shown to correlate with cell volume because of a higher overall transcription in larger cells [330, 331]. Although brain neurons possess a highly active transcriptome, the volume of their cellular bodies is reduced compared to that of hepatocytes, ranging from 10 to 15 µM in the case of paraventricular oxytocin-producing neurons [503]. Therefore, a lower number of detected transcripts and genes from brain nuclei could be expected compared to liver nuclei (**Figure 16**), because in the liver, cell size increases with the ploidy status and it has been shown in mouse and human liver that the nuclei volume of hepatocytes approximately doubles with the doubling of the DNA content [95, 328, 329].
In this study, a mean of 2,437 genes per nucleus was detected, positioning within the range of other snRNA-seq studies reported in brain tissue, such as the study published by Hu *et al.* using sNucDrop-seq in which they detect 1,662 genes per nucleus [473], the 1,791 genes per nucleus detected by Thrupp *et al.* [333], or the 2,419 transcripts per nucleus detected in Dowsett *et al.* [504]. Habib *et al.* using DroNc-seq detected a higher average of genes per nucleus with 2,731, however from human samples [26].

Importantly, the application of the snRNA-seq2 method allowed for the detection of the genes *Aldoc, Apoc1, Fth1, Rpl29, S100b, ApoE, Lepr,* and *Cst3* that could not be detected in other nuclear datasets

from human hypothalamic samples (**Figure 19**) [333]. This is highly relevant since these genes are required for microglial activation and can be used to measure response to amyloid plaque accumulation in mice, as a signature feature of Alzheimer's disease [505]. In addition, a small fraction of nuclei was characterized by the simultaneous expression of oxytocin (*Oxt)* and prohormone pro-opiomelanocortin (*Pomc),* as well as marker genes of other neuronal subtypes (**Figure 18**). These results highlight the superior sensitivity of this methodology for the interrogation of rare cell type populations from previously archived frozen samples of complex tissues beyond the liver, such as the murine brain. Therefore, the optimization of the nuclei isolation from the neurons and the subsequent nuclei lysis steps in the microtiter wells have been essential for the identification and characterization of the gene expression of murine oxytocin-positive hypothalamic neurons. This methodology allowed for the detection of a high number of transcripts per nuclei, identifying genes that were found depleted in other brain datasets from human samples, thereby highlighting the high sensitivity achieved with this method.

4.2 Investigation of the heterogeneity in the liver using single-cell genomic approaches

4.2.1 Low-throughput scATAC-seq plate-based approach to explore polyploidy at the epigenomic level

In this chapter, the aim was to enable the assessment of the heterogeneity in the chromatin accessibility configurations in polyploid hepatocytes (2n and 4n) from murine liver at single-cell resolution by optimizing a low throughput, plate-based scATAC-seq approach. In the liver, the exploration of the epigenetic landscape in a complex starred feature of the tissue, such as liver polyploidy has to date not been addressed. Here, the initial steps for the development of a protocol that allows for the study of complex tissue characteristics requiring a specific experimental design such as liver polyploidization using a plate-based approach are laid.

Due to the high throughput, cost, flexibility, and capture efficiency, the commercialized high-throughput droplet-based scATAC-seq kit from 10X Genomics has become the method of choice for studying genome-wide chromatin accessibility [63]. Nevertheless, as discussed in the previous section of this thesis, plate-based methodologies present advantages for certain applications. Plate-based methodologies for scATAC-seq mostly rely on the upfront tagmentation of nuclei in a

bulk reaction followed by FACS sorting the nuclei into 384-well plates [60, 65, 66]. However, sorting nuclei after tagmentation does not allow for the differentiation of different ploidy levels (2n and 4n) in nuclei (**Figure 21**). Therefore, the optimization of a scATAC-seq method on plates that allows for the deep characterization of ploidy levels is necessary.

In order to achieve that, I combined the method published by Chen *et al.* in 2018, where 384-well microtiter plates and an upfront bulk tagmentation is used [60], and the method published by Buentrostro *et al.* using the ICELL8 microwell system from Takara [44]. The latter provides the advantage of a more efficient tagmentation reaction occurring per single well after cells have been distributed onto the microfluidics system [44]. However, it requires access to a specialized device and expensive consumables, as well as a laborious protocol. On the other hand, the protocol published by Chen *et al.* provides the advantage of using commonly obtainable 384-well microtiter plates and a FACS sorting system. Additionally, the use of digitonin in the tagmentation buffer provides stronger lysis of the cellular membrane. The addition of a Tagmentation Stop Buffer stops the tagmentation reaction, ensuring adequate tagmentation time per single well [60]. Combining the two protocols enabled the initial steps for the exploration of liver ploidy levels by sorting the nuclei before the tagmentation reaction.

The combination of these two methodologies, namely the use of commercially available reagents and commonly used consumables and machinery allows to reduce the cost per run and the need for specialized equipment. This enables the reduction of experimental costs by prescinding from specialized machinery and improving the method further would facilitate the interrogation of convoluted biological questions tailored to specific tissue characteristics.

In the liver, polyploidy is an example of a key feature that introduces heterogeneity among hepatocytes, however, the heterogeneity present in the liver is also impacted by other hepatic functions like the metabolism of carbohydrates, proteins, and fat, as well as xenobiotic substances such as drugs, playing an essential role in the general regulation of the energy homeostasis in the body and detoxification of xenobiotic substances [69-72].

Overall, there are several factors that affect the heterogeneity within the liver tissue. Internal factors are comprised for instance, of the presence of polyploidy [506], but also fat accumulation in the tissue, as occurs in NAFLD [507]; the presence of SNPs affecting key metabolic enzymes [117]; or aging [170, 508]. The impact of these factors on the metabolic function of PHHs was explored in this thesis. The heterogeneity present in a population of PHHs in vitro was investigated using single-cell genomics approaches at the transcriptomic and epigenomic levels. Thereafter, the capacity of individual hepatocytes to metabolize a cocktail of drugs was assessed using a phenotyping drug cocktail.

Moreover, the study of the impact of intracellular fat accumulation as occurring in NAFLD on the metabolism of *in vitro* cultured PHHs was performed. In addition, how lipid accumulation impacted the drug-related metabolism of individual hepatocytes was explored at single-cell resolution.

4.2.2 *In vitro* models for the study of hepatic drug and lipid metabolism

Several *in vitro* systems have been developed for toxicological studies, disease or drug development, including PHHs, liver hepatoma cell lines and pluripotent stem cells [232, 238, 252, 509-511]. Immortalized hepatoma cell lines such as HepG2 or HepaRG provide the advantage of being sustainable in time due to clonal expansion yet retaining certain characteristics comparable to liver tissue or PHHs [512]. For intracellular fat accumulation studies, the HepG2 and HepaRG cell lines are a suitable model, reaching similar levels of lipid storage to those detected in FFA-overloaded PHHs [260, 367, 513-515]. A factor studied in this project was the chronic intracellular fat accumulation in PHHs *in vitro*, mimicking the phenotype observed in NAFLD. *In vitro* models of NAFLD have been developed to characterize the mechanisms of its development and progression [264, 270, 516], which hallmark is lipid accumulation, as a result of a dysregulated lipid metabolism, increased lipogenesis, and reduced lipolysis [517, 518]. Intracellular lipid accumulation *in vitro* can be induced by incubating the cells with FFA, the most extensively used being oleic and palmitic acid. The optimization of susceptible steps *in vitro*, such as the incubation time, the FFA ratio, and the concentration of FFA leads to different phenotypic lipid storage and lipotoxicity outcomes, triggering apoptosis and cell death depending on high concentrations or prolonged *in vitro* incubation times [260, 367, 513-515]. Indeed, the incubation of HepG2 with a 2:1 ratio of oleic to palmitic FFAs in three different concentrations (100 µM, 150 µM and 200 µM), led to intracellular fat accumulation confirmed by Oil-Red-O staining of lipids, without causing cell detachment or proliferation arrest (**Figure 24**). This concentration mimics dietary conditions observed *in vivo* inducing NAFLD [260]. Therefore, these conditions were selected for the experiments performed in HepaRG cells in bulk, and in PHHs at single-cell resolution. Optimization of susceptible steps *in vitro*, such as the incubation time, the FFA ratio and concentration or the obtention of a single-cell suspension, as well as a curated experimental design prior to the performance of technically and logistically demanding single-cell experiments, are key for yielding a successful experiment and a high-quality dataset.

To assess the metabolic capacity of the cells, the strategy conventionally known as the “cocktail approach” [272, 281-285, 519] was used. However, this strategy has not been extensively used to

phenotype and monitor the metabolic capacity of the liver by assessing the enzymatic activity of the CYPs by observing changes in the induction of their mRNA level in bulk and the protein level. Typically, the use of phenotyping probe cocktails serves for the investigation of drug-drug interactions (DDIs) and the pharmacokinetic assessment of CYP450 enzymes focusing on the measurement of plasma or urine concentrations of metabolites or mass spectrometry or chromatography metabolite analysis [277, 285, 520, 521], and infrequently, measuring mRNA level changes upon cocktail incubation *in vitro* [287, 522, 523].

The selection of the Sanofi-Aventis phenotyping probe cocktail among the several available ones was due to its successful usage with high specificity and safety in human studies [285, 524] and other primates [305, 525], mouse, dogs and minipigs [304, 526] in both *in vitro* [304] and *in vivo* [285, 287, 305]. The five probes composing the five-drug cocktail are substrates for the main isoforms of the CYP450 monooxygenase enzymes: *CYP1A2, CYP2C9, CYP2C19, CYP3A4*, which are responsible for the metabolism of approximately 70-80% of the nowadays commercially available drugs [117, 123, 527, 528]. *In vitro,* the drug catabolism is defined in three phases, and phase I refers to oxidation, hydrolysis, reduction, and cyclization reactions that are catabolized mainly by the CYPs. The expression of these CYPs is induced or inhibited by their substrate compounds, which are generally used as a measurement for the functional characterization of the liver metabolic phenotype. However, there are post-translational mechanisms involved in the correlation of gene expression level to the protein or enzymatic activity. A tight correlation between mRNA and protein levels of most of the CYP isoforms has been described in bulk and at subpopulation resolution [123, 259]. For instance, it has been reported that the expression of *CYP1A1, CYP1A2, CYP3A4, CYP2D6,* and *CYP2B6* is mostly determined by pre-translational regulatory mechanisms, and estimating the CYP mRNA level seems appropriate to infer CYP protein activity [123].

Prior to its application on PHHs, the functionality of the drug cocktail on the CYP450s expression was assessed (Sanofi-Aventis) [285] in HepaRG cells by qPCR (**Figure 26**). A similar approach using different probes to assess CYP450 mRNA levels was used by applying the Basel cocktail by Berger *et al*. in HepG2 cells, HepaRG and cryopreserved PHHs in 2D and 3D culture[137]. This probe cocktail differs from the Sanofi Aventis cocktail in the use of losartan instead of S-warfarin as a substrate of CYP2C9 [137]. However, the broadly used fold change calculation using the $2^{-\Delta\Delta CT}$ method for qPCR measurements relies on the assumption that the efficiency of amplification is 100%, and the same in all samples and the amount of molecules doubles with each PCR cycle, obtaining an efficiency of 2. This has been empirically proven to not always be the case [529]. This improved method takes into consideration the PCR efficiency of each individual sample and removes the background fluorescence,

resulting in higher accuracy in the calculation of the relative gene expression [529]. Therefore, in this study, the improved $2^{-\Delta\Delta CT}$ method published by *Rao et al.* was used to calculate the precise amplification efficiency per sample and per target gene amplified in the qPCR performed on HepaRG (**Figure 26**). This resulted in a more accurate calculation of the amplification reaction efficiency for each of the targets in the qPCR, for instance, the efficiency calculated for the housekeeping gene *PGBD* was lower than 2, whereas that of *CYP2C9* was higher than 2 and that of *CYP2C19* was 2. An efficiency lower than two represents a PCR reaction efficiency lower than 100%, where the molecules of cDNA do not duplicate with each PCR cycle [530, 531]. However, a higher efficiency than 2 has been shown to be due to overconcentrated cDNA template, the presence of primer dimers, or PCR reaction inhibitors [370, 532]. The accurate calculation of the PCR reaction efficiency is of greater importance when using intercalating DNA dyes such as SYBR Green.

In the *in vitro* model used here of differentiated HepaRG cells, an upregulation of the *CYP2C9*, *CYP2C19,* and *CYP3A4* expression (mRNA level) upon Cocktail treatment was observed, detecting a 0.55-fold, 0.48-fold, and 3.65-fold induction, respectively (**Figure 26**). This confirmed the competency of this cell line for the investigation of drug metabolism and CYP induction studies, as well as the probes in the cocktail being able to induce the gene expression of the CYP enzyme functionality [240, 287]. Moreover, it has been previously shown that in human livers, the correlation of the mRNA level measured by RT-PCR and the respective protein level hold a tight correlation for the CYP isoforms *CYP1A1, CYP1A2, CYP3A4,* and *CYP2D6* [123]. Using the Basel cocktail and rifampicin incubation for CYP pre-induction, the mRNA level upregulation of *CYP1A2* and *CYP2D6* using caffeine and metoprolol as probe substrates was not observable [137]. However, it was observable for *CYP2C9*, *CYP2C19* and *CYP3A4* [137].

Upon free-fatty acid incubation, a downregulation of the expression level of *CYP2C9, CYP2C19,* and *CYP3A4* compared to DMSO level was measured, and the largest magnitude change was observed for *CYP3A4*, showing a 0.62-fold downregulation (**Figure 26**). This is in concordance with previous findings in lipid-loaded HepaRG cells, reporting a downregulation of *CYP3A4* mRNA levels [533, 534]. When lipid-loaded HepaRG cells were exposed to a five-drug cocktail incubation, intracellular fat accumulation hindered the upregulation of the CYP450 enzyme levels. Only *CYP3A4* showed a 0.32-fold upregulation over the DMSO baseline level. Controversial reports on the effect of hepatic fat accumulation on the CYP450 enzyme expression and activity have been published [535], and there is not a clear consensus on the effect of fat accumulation on drug metabolic capacity in HepaRG cells. Furthermore, known disadvantages of HepaRG cells are that this cell line is derived from a donor that had a poor metabolizer allele for CYP2D6, CYP3A5, and CYP2C9 [536]. This poses a disadvantage because CYP2D6 is, together with CYP3A4, a major CYP450 enzyme, metabolizing around 30% of

the drugs in the market. However, HepG2 cells do not retain a high expression of key liver transcription factors such as HNF4A, CAR or PXR regulating the activity of CYP450 xenobiotic-metabolizing enzymes [234-236, 537]. This deems them not suitable for CYP450 enzyme induction studies related to drug metabolism, disposition, and drug toxicity [240-242]. For this purpose, the hepatoma cell line HepaRG represents a closer model of PHHs physiology and metabolic capacity [231, 233, 238, 538]. The investigation of toxicology, safety and efficacy, and drug-drug interactions during drug discovery and development occurs in *in vitro* models of the liver prior to *in vivo* phases [427, 539]. The gold standard *in vitro* model to study drug metabolism and toxicity are PHHs isolated from resected liver tissue because they provide a close representation of the metabolism and functionality of the liver in humans [229, 540, 541], sharing morphological, biochemical, and resemblance to *in vivo* hepatocytes[373, 541-543]. Moreover, they retain the natural capacity to induce CYP450 enzymes, but they also present limitations such as the inter-donor variability, the lack of proliferation capacity [544], and their dedifferentiation losing the hepatic phenotype and gene expression levels when cultured in monolayer for days [253, 545-547]. Despite these, cryopreserved PHHs present, they are an advantageous tool for the assessment of drug metabolism and hepatotoxicity of newly developed drugs [548-550].

In this thesis, the capacity of individual PHHs to metabolize drugs, as well as the effect of intracellular lipid accumulation in their metabolic profile was evaluated. In addition, how intracellular lipid accumulation impacted the capacity of single PHHs to metabolize drugs was assessed both at the epigenomic and transcriptomic levels.

4.2.3 Exploration of the chromatin accessibility configuration in primary human hepatocytes

Different euchromatin configuration statuses dictate the downstream gene expression profile by regulating the access of transcription factors to binding sites in enhancers and promoters [551], which have been associated with different conditions and diseases in several tissues [45, 551, 552]. Single-cell ATAC-seq is a widely used methodology to investigate the transposase-accessible chromatin heterogeneity comprised in tissues and cell populations [45, 56, 348]. In the liver, single-cell ATAC-seq has recently revealed heterogeneity in the chromatin accessibility pattern and TF modulation configuration among the different hepatic cell types [45] during liver regeneration [48], aging [111] and in NAFLD and NASH disease progression [215-217].

Applying a droplet-based approach allows for the study of thousands of cells simultaneously in high-throughput [47, 553, 554], which is of advantage in the identification and characterization of unique cell states within the same cell type, and or individual cellular responses on a single-cell level [555].

However, the physical properties of the tissue or cell type of interest determine several steps in the protocol that need to be optimized and adjusted to obtain a successful experiment [68]. In the present thesis, the steps prior to the performance of scATAC-seq using 10X Genomics Chromium technology were optimized for its application on PHHs. At first, an experiment using 10 min or 30 min incubation in lysis buffer was designed, allowing for the selection of the 10 min lysis incubation time for subsequent experiments based on the quality control performed including TSS enrichment and fragment length distribution, although similar electropherogram profiles were obtained between samples (**Figure 28** and **Figure 29**) [556]. Overall, this sample showed a higher quality and the TSS distribution calculated showed less noise vs. signal ratio. However, after microscopical inspection, the cellular membrane of some PHHs was not fully dissociated. Therefore, the combination of the methodology using a homogenization buffer supplemented with 0.3% Triton X and 0.2% NP-40 [19], a swelling buffer [48] and mechanical dissociation using a Dounce homogenizer proved to be successful in dissociating the robust cellular membrane composing PHHs after *in vitro* culture [375]. It is worth considering that sample handling becomes technically and logistically challenging when a higher number of samples are processed simultaneously, especially when the experimental protocol is laborious and is composed of several critical time-dependent incubation steps. On the other hand, performing the experiment in two different batches might introduce technical bias, which might increase the need for a stricter batch correction in the downstream computational analysis [557-560]. In the experiment performed on eight PHHs samples at single-cell resolution in the four experimental conditions (DMSO, Cocktail, FFA+DMSo and FFA+Ccocktail), tagmentation excess was observable in six out of eight samples in the experiment. Therefore, this could be taken into consideration for further replicates, drawing special attention to lysis buffer incubation time. Nevertheless, the experiment performed here yielded scATAC-seq data for preliminary downstream biological investigation. The preliminary scATAC-seq data presented in this thesis is constrained to two human liver donors (Methods).

A challenge when analyzing and working with scATAC-seq datasets is the selection of the feature to build the count matrix [561]. In the present thesis, the count matrix was built on called peaks and embedded in the same feature space (Methods). Although normalization scales count data to render

peak counts comparable between cells, a count depth effect often remains in the data. In this study, the number of counts was identified as a confounding factor, leading to their regression as a normalization strategy (**Figure 35** and **Figure 36**). Linear regression can be applied to subtract technical variation, and the most prominent technical covariates in single-cell data are count depth and batch [562]. This counts depth effect can be both a biological and a technical artifact, most likely the latter in this dataset. Therefore, linear regression was applied here as a normalization technique after confirming that the number of counts correlated to PC1 (R^2=0.99) [376, 562]. Yet, technical count effects may remain after normalization as no scaling method can infer the expression values of genes that were not detected due to low-quality samples.

To explore the heterogeneity in the chromatin accessibility configurations between the four experimental conditions used on PHHs *in vitro* (DMSO, Cocktail, FFA+DMSO and FFA+Cocktail), differentially accessible peaks were called and ranked using *scanpy* (Methods, **Figure 37**). Among these, the differentially accessible regions upon cocktail treatment correspond to intergenic regions. These regions in fact constitute the majority of the genome and are non-protein-coding, typically composed of enhancers and promoters [375, 379, 380], and these regions regulate gene expression [563]. Peaks were annotated by genes 5 kb downstream but the most differentially accessible in cocktail treatment were not found near a gene. These numerous intergenic regions were detected because they compose the majority of the genome, they are therefore mostly captured in the dataset, and they are regions differentially accessible between treatment conditions.
Differentially accessible peaks were detected in each of the four conditions in the study (**Figure 38**). Upon vehicle (DMSO) treatment, peaks in proximity to genes responsible for cellular identity and essential cellular function maintenance were detected, for instance, *ALB* [564], *NAPA,* or *ATF3* [565]. However, upon cocktail treatment, peaks in proximity to phase I enzymes were detected, such as *CYP2C9*, *CYP3A5*, and *CYP2B6,* or phase II enzymes such as *GSTO1* (**Figure 38**). This suggests that exposure to a metabolic challenge such as a five-drug cocktail leads to a change in the chromatin accessibility pattern in primary hepatocytes, opening promoters of specific drug-metabolizing enzyme genes.
Interestingly, a peak in proximity to the liver-enriched transcription factor *HNF4A* was differentially accessible in cocktail-treated cells (**Figure 38**). This indicates that hepatocytes treated with a five-drug cocktail, or free fatty acids could synthesize more HNF4A, which could then lead to increased drug- and fat-related metabolism if HNF4A itself binds to promoter regions of the genes it regulates. In the adult human liver, this transcription factor regulates key genes such as albumin and others involved in glucose, lipids, cholesterol, and drug-related metabolism [71, 566, 567].

When PHHs were presented with simultaneous lipid accumulation and a five-drug cocktail, peaks in the proximity of genes related to eicosanoid and drug metabolism such as *CYP4F11* or *CYP2J2* [568, 569] were detected. Overexpression of *CYP2J2* in mice has been found to attenuate high-fat diet-induced NAFLD [570]. Upon FFA+Cocktail treatment, peaks upstream of genes involved in inflammatory processes such as *CYP4F11* [390] and the amino acid transporter *SLC7A2*, which has been previously related to inflammation and immunity diseases such as hepatocellular carcinoma [394, 395] were differentially accessible. This indicates that, upon exposure to both intracellular fat accumulation and a five-drug cocktail, the chromatin accessibility arrangement of PHHs changes to open regions in proximity to genes related to inflammation and cellular transmembrane transport.

Taken together, these results indicate that environmental factors such as exposure to a metabolic challenge constituted by exposure to several drugs simultaneously or chronic intracellular lipid accumulation differentially alter the chromatin accessibility configuration of PHHs. Moreover, the concomitant occurrence of fat accumulation and exposure to a cocktail of drugs triggered a different chromatin openness response. Hence, it can be postulated that these epigenomic changes toward environmental conditions such as diet, disease, or drug exposure can modulate the downstream gene expression patterns observable at the transcriptomic level.

4.2.4 Heterogeneous transcriptional responses among primary human hepatocytes and influencing factors

During drug discovery and early drug development phases, safety, toxicity, and efficacy are generally tested on PHHs in bulk analyses, thereby assuming homogeneity among hepatocytes and obtaining an averaged readout that represents the traits of the most abundant cell subtype [239, 253, 260, 571, 572]. However, this approach hinders the assessment of cellular heterogeneity the identification and characterization of cellular subpopulations with differential phenotypes and transcriptional responses towards a stimulus [573]. In the present thesis, the heterogeneity among PHHs *in vitro* has been revealed using a high-throughput droplet-based approach (10X Genomics), leading to the identification of hepatocyte subgroups within a seemingly homogeneous cell population from a complex tissue like the liver. For this purpose, the application of a droplet-based method is advantageous, having a relatively low cost per cell, and maximizing the throughput of cells [28].

Here, four subgroups of hepatocytes were identified after 72 h of *in vitro* culture, independently of donor and treatment condition (**Figure 42**). In concordance with previous research on PHHs *in vitro*, a subgroup of cells (subgroup IV) losing their mature hepatocyte transcriptional signature was identified after 72 h in culture by calculating the percentage of cells in which a gene is expressed per subgroup. This subgroup showed the lowest overall gene expression, among which the CYPs were present (**Figure 44**) [239, 253, 425, 426, 574]. In this subgroup, a downregulation liver-enriched transcription factors that modulate the mature hepatocyte phenotype including *MLXIPL* (*ChREBP*), *RXRA*, *NH1H4* (*FXR*), *PPARA*, *HNF4A,* and *CEBPA* was described [1, 575-577]. This is of relevance in preclinical phase pharmaco- and toxicological studies classically carried out in bulk analysis using PHHs *in vitro.* Bulk analyses represent an averaged readout of the cell population [571, 572], potentially leading to the underestimation of drug efficacy and safety caused by the results being masked by cells losing mature hepatocyte gene expression [278]. Furthermore, the main reason for drug attrition in the early phases of drug development and discovery *in vitro* processes is the absence of clinical efficiency [578] or low efficacy [579], and the absence of the desired drug effect and toxicity on the cells [580-583]. If a subgroup of cells in the culture does not behave anymore as mature hepatocytes due to *in vitro* phenotypic instability and shows a comparatively reduced metabolic capacity, the overall obtained readout assessment is biased by the influence of these cells.

Once the presence of subgroups in the PHH *in vitro* population was unravelled, a metabolic challenge was posed by a cocktail of five drugs (Methods). Upon exposure to the Sanofi-Aventis five-drug cocktail for 66 hours, the four hepatocyte subgroups showed distinct transcriptional responses to this metabolic challenge, which were not distinguishable in pseudobulk analyses (**Figure 45**). A clear example is the 4-fold upregulation of *CYP3A4* found in subgroup I under Cocktail treatment compared to DMSO level, whereas no significant upregulation was detected in subgroup III (**Figure 45**). This indicates that differences in the single-cell gene expression profiles of hepatocyte subgroups impact the mean and variance of bulk or pseudobulk gene expression readouts [584]. Moreover, these metabolic differences between subpopulations are concealed unless performing single-cell readouts.

Furthermore, the subgroups showed specialization towards different metabolic pathways regarding both endo- and xenobiotic substances. This analysis was performed through GO analysis using a database suited for toxicogenomic comparisons (CTD). The results revealed a targeted functionality regarding the xenobiotics that the different subgroups can metabolize. For instance, subgroup I showed overlaps with the metabolic routes for the metabolism of crizotinib, which is a compound used for non-small cell lung cancer treatment [585]. This also shows the susceptibility of each subgroup to

develop toxic metabolites leading to adverse reactions when exposed to a certain compound (**Figure 45**). It would therefore be crucial to consider that the here identified hepatocyte subgroups show distinct metabolic capacity and transcriptomic responses to exposure to different compounds, and especially toward a metabolic challenge like a drug cocktail. This is of importance when assessing drug safety, efficacy, and functionality using PHHs in culture in bulk analyses, due to the possibly confounded averaged readout. Especially because CYP induction and inhibition cause potential drug-drug interactions, hepatotoxicity, and adverse drug reactions (ADRs) such as drug-induced liver injury (DILI), potentially leading to acute liver failure, and this effect cannot be anticipated [586, 587]. In addition, ADRs are the main cause of direct drug attrition during drug discovery and development and drug market withdrawal [278, 582, 588-592]. Hence, a detailed understanding of the heterogeneity among cellular responses to drug administration at the cellular level as shown here would be beneficial towards the prediction of liver damage at an early stage. The identification of hepatocyte subgroups and their proportions among the population of cells in culture before drug candidate testing, accounting for the presence of heterogeneity in the cellular metabolic responses towards a drug, would enable a more accurate test readout and the prediction of hepatotoxicity and liver damage at an early developmental stage.

a) Pharmacogenomics: Genetic variation and single-nucleotide polymorphisms (SNPs)

Another factor affecting the inter-individual drug-related metabolic capacity observed *in vivo* may have a genetic, epigenetic, or environmental cause. Among genetic variation, the presence of single-nucleotide polymorphisms (SNPs) has been described as a major factor responsible for the variation observed in different individuals and ethnicities [117]. Moreover, they constitute a known genetic source of increased risk for the development of ADRs, for instance when they affect xenobiotic clearance pathways, such as phase I enzymes (CYP450s) [117, 593, 594].

The gap between the genome and gene expression is composed of a sequential chain of reactions at different levels. Namely, gene expression depends on the genetic information, the methylome, and in general, epigenetics and the chromatin architecture [595, 596]. To perform SNP detection at single-cell resolution, it would be ideal to use scDNA-seq data, however, the scDNA-seq methodology is not extensively used and it is a laborious method, and easily applicable commercial kits and computational pipelines for analysis are scarce [597, 598]. When using scRNA-seq data to identify SNPs, full-length transcript data would be preferable [599]. However, 10X Genomics has developed the Long Ranger

tool to infer SNPs from 3' captured mRNA data [412]. Long Ranger allows to infer single-nucleotide variations from transcriptomics data and the variants called here were searched for in the gene bodies, excluding those located in enhancers or promoters of the genes and transcription factor binding sites.

The SNP in the *CYP2C19* gene (*CYP2C19*4*, I331V) was detected in both analyzed donors (Caucasian and Hispanic). Recently, the *CYP2C19*4B* allele was discovered to have important implications for clinical *CYP2C19* testing as the allele harbors both gain-of-function [c.−806C>T (**17*)] and loss-of-function [c.1A>G (**4*)] variants on the same haplotype [417, 418]. The *CYP2C19*4B* allele was identified in both Caucasian and Hispanic populations in a low frequency (≤1%)[319, 417, 418, 600-602]. A lower expression of *CYP2C19* was observed in the Hispanic donor compared to the Caucasian donor, but a generally lower expression of all CYPs was also detected (**Figure 40**). In the analysis of the dataset of two donors, the presence of missense variants in the gene body of the *CYP2C19* enzyme could lead to lower enzymatic activity to clear the drug omeprazole, administered in the phenotyping five-drug cocktail as *CYP2C19* substrate.

Depending on the CYP-specific oxidation capacity, different terms are used to refer to the associated pharmacokinetic phenotypes. In the case of polymorphisms in the *CYP2D6* and *CYP2C19* isoforms, the population is divided into poor (PM), intermediate (IM), extensive (EM), and ultrarapid metabolizers (UM) [117, 152-155]. UMs possess gain-of-function variants; EMs are the most commonly found phenotype in the population; IMs carry one functional and one functionally deficient allele resulting in reduced drug oxidation capacity; and PMs are carriers of the null allele with complete lack of function. *In vivo*, this has major implications for the individuals carrying the missense variant (PMs or IMs), needing to receive an adjusted and appropriate dose of the drugs mostly cleared by the enzyme where the variant is detected [117, 417].

In addition, distinct frequencies in the SNPs can be found in different ethnicities [117]. The *CYP2C19* mutation (*CYP2C19*4*, rs3758581, I331V) was detected in both donors (**Table 7**). The alternative allele frequency among different ethnicities such as African, American, Hispanic, East and South Asian, and European range between 89% (South Asian) and 100% (African) [419]. Therefore, the likelihood of detecting this mutation in all backgrounds is rather high. However, for other variants, the distribution of the presence of the variant changes drastically in different ethnicities. This is a factor of major importance *in vivo,* but it also has implications to consider in drug testing studies in preclinical *in vitro* phases. Drug candidates are classically tested for hepatotoxicity in preclinical phases using PHHs cultured in a monolayer proceeding from a single donor [373, 541-543]. Therefore, depending on the presence of a SNP in a *CYP450* isoform metabolizing the tested compound leading to a lower enzymatic activity, which is often associated with the ethnicity of the donor, could affect the drug

metabolism, can lead to toxicity and thereby to the incorrect assessment of the drug's efficacy and safety.

A second example is the polymorphism in the *CYP2C9* gene (*CYP2C9*2*, rs1799853) detected in the Caucasian donor (**Table 7**) representing a missense variant [319, 413, 414]. The mRNA level of *CYP2C9* detected in the Caucasian donor was higher compared to that observed in the Hispanic donor not carrying the missense variant (**Figure 40**). However, it is known that the posttranslational mechanisms dictating translation rate, protein half-life, and the presence of SNPs can perturb a correlation between mRNA level and enzyme activity [123]. In order to estimate an accurate readout on the phenotypic effect of the SNP, the enzymatic activity would need to be measured and compared to the mRNA level [603]. Dysfunctional *CYP2C9* enzymatic activity has been linked to this mutation [415], resulting in diminished drug clearance by those individuals carrying the missense variant. Individuals possessing this mutation present, for instance, a higher risk of bleeding when administered anticoagulants like warfarin [604], phenprocoumon [605], or acenocoumarol [117, 606]. In addition, it entails lower drug metabolism capacity and higher risk for developing adverse events such as hypoglycemia when administered sulfonylureas hypoglycemic drugs [607], and non-steroid anti-inflammatory drugs (NSAIDs) [608]. This highlights the importance of the detection of these mutations and their crucial high clinical impact.

Pharmacogenomics, and particularly the presence of SNPs is a factor to consider in precision medicine to develop drug therapies tailored to the genetic and inter-individual variability intrinsic in the human populations. For instance, poor or intermediate metabolizers may need a lower dose of the same drug to achieve a certain therapeutic effect and avoid undesired toxic or off-target effects. On the other hand, ultrarapid metabolizers may need a higher dose to reach an equivalent therapeutic effect. Further developments in precision medicine and clinical trials are necessary to achieve tailored therapies that maximize therapeutical success and minimize toxic or adverse reactions. Moreover, it is crucial to take into consideration the presence of SNPs in key drug-metabolizing enzymes and their specific distribution among the population of determined ethnicities, especially those in which the frequency of a variant leading to impaired or lacking drug oxidation capacity is high. This has further implications in preclinical drug testing phases, specifically in assays performed using PHHs obtained from a single donor, who could be a carrier of a variant conferring poor drug-metabolizing capacity, leading to misleading drug efficacy or safety readouts.

b) Aging and drug metabolism

A further factor affecting the inter-individual variation and CYP450 enzyme's expression and activity impacting the drug oxidation in the liver is aging [609]. Specifically in the liver, aging distresses hepatocyte function [454, 609], causing dysregulation in some CYP450s enzyme activity and expression, perturbing the drug oxidation capacity [610, 611], and directly affecting therapeutic efficacy and safety in the older population [208, 408, 455-457].

The literature on the effect of aging on phase I enzymes expression and activity remains thus far ambiguous, reporting controversial results regarding up or downregulation of CYP levels [175, 408]. In the preliminary results shown in this thesis, old cells showed a decreased expression of the main five CYP450 enzymes, albeit retaining the subgroups' differential expression profiles (**Figure 54**). Conversely, higher expression levels of *CYP2C9*, *CYP1A2,* and *CYP3A4* in old compared to young cells were observed upon simultaneous fat accumulation and five-drug cocktail treatments, and to only Cocktail treatment within cells of the same the same age (**Figure 54**). Despite the donor sample size being too small for performing statistical tests, these preliminary results suggest that, presumably, aged PHHs *in vivo* could show distinct transcriptional responses compared to young towards xenobiotic oxidation, especially when exposed to fat accumulation, or when lipid-loaded cells are exposed to a five-drug cocktail. This is of major importance because the coadministration of drugs, also known as polypharmacy, in the elderly to treat age-related comorbidities in the elderly population is a common clinical practice [168, 612-614], leading to a higher risk for the development of drug-drug interactions (DDIs), ADRs and more specifically DILI in the elderly, and these cannot thus far be anticipated [457, 615-617]. Therefore, considering the elderly population of donors in early *in vitro* phases of drug development and *in vivo* drug testing trials would be crucial, especially because the elderly are normally underrepresented [618, 619].

Lipid metabolism dysfunction has been demonstrated to be associated with aging [178], leading to hepatic steatosis and thereby increasing the NAFLD incidence in the elderly population [180]. Recently, cellular senescence has been postulated as the driver of age-dependent hepatic steatosis [173, 174]. In general, cellular senescence plays a crucial role in age-related phenotypic changes [620], including the secretion of proinflammatory cytokines that are part of the senescence-associated secretory phenotype (SASP) [621, 622], and increased oxidative stress due to mitochondrial dysfunction [169]. In addition, aging has been linked to stress in the liver through tissue functionality deterioration and lipid accumulation, including the reduction in functional liver mass volume, blood flow, telomere attrition, genomic instability, inflammation, mitochondrial dysfunction, and increased

oxidative respiration with higher oxidative stress and ROS levels [169, 170, 623-626]. In both NAFLD and aging, due to fat accumulation and lipid metabolism impairment, there is increased oxidative stress, ROS production and inflammation [356, 461]. At single-cell resolution in this PHHs dataset, old hepatocytes showed an upregulation of key marker genes for inflammation and stress-related response such as *GSTA1* or *LGALS1* [627, 628]. Interestingly, *LGALS1* was also detected as differentially upregulated in subgroups II and IV upon fat accumulation (**Figure 55**). In line with this, it has recently been reported that a significant downregulation of liver LBP levels promotes oxidative stress and inflammation, leading to the aggravation of NASH progression [465]. In this study, a lower expression of LBP was detected in aged vs. young PHHs. Moreover, it is known that the tripeptide glutathione (GSH) acts as a cellular protector of oxidative stress, and its synthesis is catalyzed by the glutathione synthetase enzyme (GS) [462, 629]. The expression of *GCLM*, coding for the modifier subunit of the GS responsible for the synthesis of GSH as downregulated in young vs. aged hepatocytes (**Figure 55**). Similarly, the expression of the known and oxidative stress regulator *TP53* [463, 464], which is also involved in senescence regulation was downregulated in old vs. young hepatocytes.
Taken together, these findings indicate that aging triggers the upregulation of oxidative stress and inflammation transcriptomic response in primary human hepatocytes *in vitro,* together with the downregulation of protective genes against fat accumulation-related stress. This has been previously reported in liver, [169, 170, 626], leading to the so-called "inflammageing" [571, 630-632].

Aging has also been shown to increase cell-to-cell variability in gene expression [185, 633], with a tissue-specific pattern [184, 188, 634, 635]. Concordant with these reports, aging increased cellular variability and a decline in the number of genes detected per cell in this *in vitro* model of PHHs (**Figure 50**). Therefore, an increment in transcriptional noise and hindered coordinated cellular transcriptional responses could be expected among PHHs upon aging [16, 185, 187, 188, 636, 637]. A recent study with mouse liver nuclei showed differential chromatin accessibility in young vs. aged cells, where cells clustered by age based on their chromatin accessibility patterns but not on their transcriptomic profile [111]. Yet, this study reports that age is a relevant factor for explaining transcriptional cell-to-cell variation. In this PHH dataset, the highest transcriptional variability detected among the four hepatocyte subgroups (**Figure 49**) was observed in subgroup IV, the subgroup of primary hepatocytes losing expression *in vitro,* composed of over 70% old cells (**Figure 53**). The subgroup of cells undergoing a dedifferentiation process in vitro was denominated as losing expression cells and showed an overall lower gene expression level. This was defined by calculating the number of cells in which a given gene is expressed (**Figure 43**) and showed downregulation of key liver transcription factors such as *MLXIPL* (*ChREBP*), *RXRA*, *NH1H4* (*FXR*), *PPARA*, *HNF4A,* and *CEBPA* [1, 375, 575, 576].

This indicates that aging might accelerate the loss of expression process *in vitro* model of primary hepatocytes, losing their mature hepatocyte transcriptional signature.

For the results obtained in the exploration of aging, it is important to consider that due to the low sample size (n=2 per age), this analysis lacks the statistical power to gain deeper insights into the transcriptomics effects of aging in this experimental setup. The encapsulation of the cocktail-treated cells sample from the young donor partially failed on the 10X Genomics chip; therefore, the cocktail treatment condition corresponds to one donor only in the "young" age in the present analysis. However, taken together, these preliminary data comparing 11,279 young and 26,953 old PHHs *in vitro* suggests that aging could accelerate the phenotypic instability and loss of expression process of hepatocytes *in vitro*. In addition, transcriptional variability was found significantly higher in older hepatocytes compared to young, together with higher stress and inflammation responses, indicating that aging leads to uncoordinated transcriptional responses.

An analysis of the transcriptomic profile per subgroup and per treatment could enable the identification of groups of cells showing a higher susceptibility to contribute to the development of toxic drug-related reactions in the elderly. Furthermore, a larger sample size in terms of the number of donors and cells per subgroup of hepatocytes to deeply characterize the transcriptional differences between subgroups of hepatocytes, to elucidate how aging affects the capacity of single hepatocytes and hepatocyte subgroups to oxidize drugs, leading to the development of ADRs and DILI in older people when exposed to the concomitant administration of several drugs. As mentioned above, the elderly are frequently excluded from clinical trials [618, 638-641].

c) Chronic lipid accumulation and drug metabolism

Chronic liver diseases such as non-alcoholic fatty liver disease (NAFLD) have been identified as a source of significant interindividual variation in metabolism [220, 535, 642]. A dysregulation of the CYP450 enzyme expression and activity has been described in NAFLD but little is known about the precise molecular mechanisms leading to these alterations at single-cell resolution in humans [535]. A reason for this gap is the lack of accurate preclinical models recapitulating the disease stages and hallmarks [643]. Concurrent with this, a higher occurrence of ADRs, especially DILI, has been observed in NAFLD patients [198, 209, 220, 644].

To assess the impact of chronic liver disease on cellular heterogeneity and drug-related metabolic capacity, hepatic steatosis occurring in early stages of NAFLD was mimicked by loading the cells with intracellular lipids along incubation time (Methods, **Figure 24** and **Figure 27**) [236, 260, 645]. The experimental conditions for the incubation with free-fatty acids were first established in HepaRG cells ("Method optimization" section in the Results).

A significant increase in cellular variability in two of the four subgroups (subgroup I and II), in contrast to the significant decrease in subgroup III clearly showcased the differential responses toward intracellular lipid storage among seemingly homogeneous hepatocytes (**Figure 49**). Particularly, it can be inferred that subgroup III, in charge of the metabolism of lipids and phase III enzymes, responds more tightly coordinated to chronic intracellular fat accumulation [16, 185, 187, 636], whereas transcriptional noise was increased among the individual cells in the other subgroups [637]. This indicates that chronic steatosis in this liver *in vitro* model of PHHs increased transcriptional variability in particular subgroups of hepatocytes and diminished their drug oxidation capacity.

Gene ontology (GO) analyses further confirmed that subgroup I upregulated lipopolysaccharide and chemokine metabolism-related pathways, as well as response to oxygen-containing compounds (**Figure 46**), indicating a possible increase in inflammatory processes upon fat accumulation exclusively in this subpopulation of hepatocytes. Hence, this specific subgroup of hepatocytes could be more involved in NAFLD development, driving stress responses. This is further confirmed by the overall higher expression of ER-stress and inflammation-related genes in this subgroup compared to the other two metabolically active subgroups (**Figure 48**). However, the pathways upregulated for subgroup II were rather related to the regulation of triglyceride metabolism and acylglycerol catabolism, and subgroup III pathways overlapped with lipid, monocarboxylic acid, and fatty acid-related metabolic processes (**Figure 46**). Certainly, GO analyses (ShinyGO) performed using the differentially upregulated genes upon FFA treatment revealed the three metabolically active subgroups (I, II and III) engage a different set of genes that are involved in distinct and specific biological pathways upon fat accumulation (**Figure 46**). This denotes that different subpopulation compositions in the liver *in vivo* could lead to diverse functional outputs and to the development of adverse events and environmental and dietary factors. In turn, this would lead to physiological conditions and diseases, in this case, chronic hepatic steatosis or NAFLD.

Due to the fat accumulation and lipid metabolism impairment, it has been shown that there is an increase in oxidative stress, ROS production, and inflammation during NAFLD [356, 461]. Also, increased oxidative stress and damage have been reported in cases of DILI, owing to an increment in

the reactive oxygen species (ROS) production [645]. In the present data, the *CYP4A11,* a CYP450 enzyme involved in NAFLD progression by inducing ROS-related lipid peroxidation and inflammation [646], was found up-regulated in subgroups I and III upon fat accumulation (**Figure 46**). The master regulator of oxidative stress *NFE2L2* also showed a higher mean expression in subgroup I upon fat accumulation (**Figure 48**). However, FGF21 which has been reported to have a protective effect against hepatic lipotoxicity [449] was significantly upregulated in subgroup III upon FFA treatment. Additionally, increased chemokine production has been observed associated with the inflammation process occurring in NAFLD [359, 448]. Chemokines *CXCL1, CXCL8, CXCL10,* and *CXCL11* were up-regulated in FFA treatment in subgroup I. Together with this, endoplasmic reticulum (ER) stress has been pointed out as a potential cause of NAFLD [446, 447]. ER stress response triggers the unfolded protein response (UPR), controlled by marker genes *ATF6* [647]*, ERN1,* and *EIF2AK3* [648]. These, together with *GDF15* [649]*, PHLDA3* [650] , and *PPP1R15A* were found more highly upregulated in subgroup I compared to the other two metabolically active subgroups (**Figure 48**). Similarly, the anti-inflammatory deubiquitinase *TNFAIP3* [651] recently shown to inactivate hepatic *ASK1* ameliorating NAFLD and protecting against disease progression [440, 441] was detected among the five top DEG genes in subgroup I of PHHs and significantly upregulated in subgroups I and III upon fat accumulation (**Figure 48**). Additionally, *ATF4*, a stress-responsive transcription factor [652, 653] found to ameliorate alcoholic liver steatosis in mice [654], was highly expressed in subgroup I compared to the other subgroups. However, subgroup I also showed a higher mean expression of *GADD45A* upon fat accumulation, which has been shown to have antioxidant properties in the liver [655], showing the activation of protective mechanisms against raised oxidative stress.
In sum, the specific response of subgroup I toward fat accumulation could mean that this specific cell population is more susceptible to stress and inflammation, while concomitantly activating protective machinery against lipid-induced damage. This highlights that the proportion of the different subgroups may play a crucial role in the assessment of functional readouts in response to dietary or environmental factors.

Regarding xenobiotic and endobiotic substance metabolism in the liver, it is known that hepatic steatosis affects the CYP450 enzyme expression and activity in hepatocytes. Studies using PHHs in bulk analyses have shown downregulation of the CYP450 mRNA levels in lipid-loaded vs. healthy cells, among them *CYP1A1*, *CYP1A2*, *CYP3A4*, *CYP2C9* and *CYP2D6* mRNA levels [225, 228]. For instance, the expression of *CYP3A4*, the major drug-metabolizing enzyme responsible for the metabolism of around 40% of the commercialized drugs, was increased in subgroups I and III upon fat accumulation vs. vehicle but maintained in the other two subgroups (II and IV) (**Figure 47**). This

indicates that there is heterogeneity in the changes in the *CYP450* enzyme expression under chronic hepatic steatosis occurring in a different manner in subgroups of the same cell type. Moreover, it suggests that different hepatocyte subpopulations possess a distinct capability to handle chronic accumulation of lipids. As such, a closer look at single-cell resolution could help our understanding of the drug oxidation capacity of subgroups of cells and altered drug metabolism during NAFLD and its progression [228]. Such studies are thus far scarce, performed in mouse models, and generally focused on other aspects of the disease [106, 210], or other disease statuses such as NASH [212].

Taking these investigations further on lipid-loaded PHHs, the phenotyping Sanofi-Aventis cocktail was used to explore the effect of the exposure to a metabolic challenge such as a cocktail of drugs. (Methods).

At single-cell resolution in the data from PHHs, Cocktail treatment induced the expression of *CYP450* enzymes in the hepatocyte subgroups (**Figure 56**), however, when the five-drug cocktail treatment was applied to cells with intracellular fat accumulation, a downregulation of the induction levels was observed for *CYP2D6, CYP2C19, CYP2C9,* and *CYP3A4* in comparison to baseline DMSO level in the four hepatocyte subgroups (**Figure 56**). This effect was larger in subgroup III, and in general, it reveals that intracellular chronic lipid accumulation diminishes the drug-related metabolic capacity of PHHs. In this case, this reduced drug oxidation capacity increases the risk of the development of xenobiotic-induced toxic reactions.

Moreover, Cocktail and FFA+Cocktail treatment only shared 234 commonly up-regulated genes, whereas 602 genes were upregulated upon FFA+Cocktail, indicating a higher involvement of biological pathways when cells were exposed to both intracellular fat accumulation and a five-drug cocktail (**Figure 57**). GO analyses revealed that these genes overlapped with stress-related pathways, indicating that the incubation with the five-drug cocktail on lipid-loaded cells showed an accentuation of the response to stress detected in FFA treatment only (without five-drug cocktail) (**Figure 57**).

Therefore, it could be postulated that lipid-loaded primary hepatocytes show a decay in their capacity to oxidize drugs upon a drug-related metabolic challenge such as exposure to a five-drug cocktail *in vitro*, distinctive for each hepatocyte subgroup. It is relevant to mention that the metabolically active hepatocyte subgroups (I, II and III) were also detected in an *in vivo* dataset with nine human donors by Aizarani *et al.* [2], which could suggest that these two factors combined (fat accumulation and exposure to drugs) could act synergistically limiting the drug metabolic capacity of liver cells in a heterogeneous manner. Moreover, hepatic lipid accumulation and healthy aging are tightly associated, which indicates that lipid accumulation may accelerate the aging process by further increasing the transcriptional variability and the inflammation and stress responses.

4.3 Cellular heterogeneity in the liver is affected by several factors leading to phenotypic differences

The work in this thesis shows that both intrinsic and external factors contribute to cellular heterogeneity and inter-donor variability in the liver. Intrinsic factors include the presence of SNPs, aging, or lipid accumulation, whereas diet and administration of drugs constitute environmental factors. For instance, inter-donor variability is represented by the presence of SNPs in key metabolic enzyme genes, affecting the capacity to metabolize drugs and the susceptibility to suffering drug-related toxicity. In addition, aging and intracellular fat accumulation increase transcriptional variability, leading to uncoordinated transcriptional responses, and activating stress and inflammation pathways. Furthermore, both aging and lipid accumulation activate stress and inflammation-related responses, which both accelerate the overall loss of expression of individual primary human hepatocytes *in vitro*, reducing their ability to metabolize a cocktail of drugs. These findings suggest that *in vivo* both aging and fat accumulation are risk factors increasing cellular heterogeneity and variability in hepatocyte subpopulations. This thesis has also shown that different of these subpopulations may be more susceptible to having aberrant responses to drugs, increasing the possibility of developing adverse drug reactions. This is of crucial importance due to the commonly concurrent hepatic steatosis or NAFLD occurring with age, producing a synergistic effect that enhances the risk for drug-related hepatotoxic outcomes.

In the comprehensive data here presented, the relevance of cellular heterogeneity in the liver can be appreciated from the transcriptomic and the epigenomic level by observing changes in the chromatin accessibility configuration that reflect the transcriptomic response toward intrinsic and extrinsic factors. These factors studied in the PHHs *in vitro* were fat accumulation, exposure to a drug cocktail, or both simultaneously. Differentially accessible regions were detected at the epigenomic level, for instance upstream of CYP450 genes and the liver-enriched transcription factor *HNF4A* upon cocktail treatment, which also showed upregulation at the transcriptomic level upon drug cocktail incubation. Intracellular fat accumulation led to differentially accessible regions upstream of specific genes related to lipid metabolism. For instance, *HNF1A*, *ABCC3*, or *APOH*, were also identified as differentially accessible and as DEGs in FFA-treated cells at the transcriptomic level, showing the modulation of the downstream gene expression by the euchromatin organization and accessibility pattern. Similarly, when lipid-loaded cells were exposed to a five-drug cocktail *SLC19A1* or *SLC7A2,* as well as the *ONECUT2* were differentially accessible, and detected as DEGs at the transcriptomic level under the same experimental condition. Therefore, a deeper characterization of the upstream molecular events

regulating downstream gene expression would shed light on the regulatory mechanisms of the functional outcomes in response to factors like fat accumulation or concomitant drug exposure.

In the future, taking into consideration the existence of cellular heterogeneity and more specifically, cellular subgroups in a seemingly homogeneous population of PHHs *in vitro* and the tissue function is crucial for drug development. In addition, the impact of SNPs, aging, and fat accumulation on this heterogeneity in *in vivo* studies is of crucial importance. In *in vitro* phases, it will lead to a more accurate drug efficacy and safety readout, especially linked to age or fat accumulation pathologies like NAFLD. Moreover, aging and intracellular fat accumulation both lead to changes in the overall subgroup proportion composing the population, and to increased transcriptional variability, indicating that the functional outcome and coordinated response may be driven by the number of cells of a certain hepatocyte subgroup. Further studies characterizing the role of the different subgroups during aging after exposure to a drug cocktail, lipid accumulation or both could help anticipate the occurrence of hepatotoxic reactions, ADRs, and DILI, which currently cannot be anticipated. Especially, because the elderly are typically excluded from drug testing trials.

5. Conclusions

In this dissertation, two different single-cell genomics approaches were used to investigate the heterogeneity intrinsic in complex tissues, specifically the brain and the liver. These two approaches diverge mainly in the throughput and sensitivity they provide, as well as in single-cell and molecular capture.

Here, the plate-based low throughput snRNA-seq2 method was applied to explore a specific neuronal population of the murine hypothalamus, resulting in the identification of subtypes of oxytocin-positive neurons showing mutually exclusive gene expression. This is represented by the expression of the neuropeptide arginine vasopressin, in the nuclei that lowly expressed *Oxt* and vice versa. This neuropeptide is known to antagonize oxytocin functions, indicating that the expression of these two is complementary and mutually exclusive, being specific to two different neuron subtypes. Overall, these results also highlight the versatile usage of the snRNA-seq2 methodology to deeply characterize rare cell populations, allowing other researchers to adapt it and use it in their research field.

The results presented in the main focus of this thesis showcase the key role of cellular heterogeneity in the comprehensive evaluation of liver functionality under different environmental factors or disease conditions. Factors affecting the heterogeneity in the liver include polyploidy, the presence of SNPs, aging, chronic lipid accumulation, or exposure to xenobiotics. Herein, the impact of these intrinsic and extrinsic factors including cellular heterogeneity was explored in this work.

Using a plate-based approach, the initial steps for the development of a methodology that allows the exploration of the heterogeneity in the chromatin accessibility configurations using scATAC-seq in complex tissue characteristics such as liver polyploidy was established. Hereby, the combination of two methodologies resulted in the obtention of the nucleosome periodicity pattern expected from scATAC-seq experiments. In the future, the method here established will allow the exploration of the impact of polyploidy in the epigenetic landscape of hepatocytes.

A high throughput, droplet-based approach was used to study the impact of the aforementioned factors on the heterogeneity present in PHHs, which are considered a homogeneous cell population. Thereby, cryopreserved PHHs from four donors were exposed to a cocktail of five drugs, loaded with intracellular lipids, or both simultaneously. Thereafter, to explore the metabolic profile of PHHs at single-cell resolution a droplet-based scRNA-seq approach was applied. At the transcriptomic level, four major subgroups of PHHs *in vitro* were identified across all four donors and independently of the treatment conditions. These subgroups showed differential metabolic profiles, specializing in the

metabolism of different endo- and xenobiotic substances. Namely, subgroup I specialized in sterol and bile acid metabolism; subgroup II in carbohydrate and phase II metabolism and subgroup III, in lipids and phase III metabolism. In concordance with previous literature, a fourth subgroup losing mature hepatocyte gene expression was identified *in vitro*. This indicates that there is cellular heterogeneity and specialization intrinsic in a seemingly homogeneous population of cells.

To assess their metabolic capacity, PHHs were exposed to a five-drug phenotyping cocktail composed of probes that are substrates of the main five CYP450 enzymes responsible for the metabolism of 70 to 80% of the commercialized drugs. The concomitant administration of drugs revealed that the hepatocyte subgroups possess different capacities to metabolize xenobiotic compounds, impacting the overall drug-related metabolic functionality of subgroups of hepatocytes in a different manner. This could be observed by the differential expression levels and cell proportions of the main five CYP between subgroups. This indicates that cellular subpopulations can have a higher susceptibility towards the development of adverse drug reactions and hepatotoxicity when exposed to a compound they cannot efficiently metabolize. Moreover, a differential chromatin accessibility configuration was observed upon cocktail exposure, where regions upstream of several CYP450 genes and the liver-enriched transcription factor HNF4A were accessible. Thereby, the cellular heterogeneity present among PHHs has an impact on the capacity of the liver to metabolize drugs. This is of high relevance because cellular heterogeneity might play a key role in the assessment of drug metabolic capacity, drug efficacy, and safety during early preclinical phases of drug discovery development.

Regarding the functionality of the liver in drug metabolism, the presence of SNPs in the CYP450 xenobiotic-metabolizing enzymes has been shown to introduce inter-individual variation and produce changes in the expression level and enzymatic activity [117, 148, 656]. From the transcriptomic data of two donors, the presence of SNPs in the main five CYP450 enzymes was inferred, and a common mutation in the *CYP2C19* gene was detected, which translated into a missense variant. One of the donors showed a lower expression level of *CYP2C19*, but also of the other four CYPs. However, a lower gene expression does not always correlate with a lower enzymatic activity. In *in vitro* phases of drug development using PHHs and in the clinical phases, the functional readout of the metabolism (efficacy), and the drug safety might be impacted by the presence of SNPs affecting the gene expression and further the enzymatic activity of key enzymes responsible for drug metabolism.

Another factor introducing variation in the metabolic profile of the liver is the presence of chronic fat accumulation or steatosis, as occurs in chronic liver diseases like NAFLD. In this *in vitro* model of

PHHs, intracellular fat accumulation was mimicked by loading the cells with free fatty acids. Lipid storage affected the gene expression in the four hepatocyte subgroups in a differential manner, showing functional specialization of the cells. Moreover, fatty acid accumulation triggered a significant increase in transcriptional variability in two subgroups of primary hepatocytes (I and II), indicating higher transcriptional noise and less coordinated transcriptional responses after exposure to fatty acids [16, 22, 185, 187, 450]. The subgroup responsible for the metabolism of lipids (subgroup III), showed a significant decrease in transcriptional variability upon fat accumulation, suggesting that this subgroup engages a robust and coordinated response to chronic intracellular lipid storage.
A further effect of fat accumulation was the increase in lipopolysaccharide and chemokine metabolism pathways specifically in subgroup I, together with increased ER-stress and inflammation markers. This indicates that chronic liver chronic disease or dietary factors associated with hepatic fat accumulation affect the proportional composition of hepatocyte subgroups and may impact the metabolic functionality of the liver. Differential chromatin accessibility regions upstream of genes related to drug metabolism were found open upon Cocktail treatment, concordant with the results observed at the transcriptomic level. Moreover, fat accumulation led to finding regions differentially open related to lipid metabolism genes. For instance, *HNF1A*, *ABCC3*, and *APOH*, which were also DEGs in lipid-loaded cells in the transcriptomics analysis. This indicated that a distinct chromatin configuration regulating gene expression impacts the downstream transcriptomics phenotypic response.

Related to this is the hepatic accumulation of fat during healthy aging, and thereby a similar phenotype could be expected. The preliminary results on aging shown in this thesis showcase a higher transcriptional variability in aged vs. young PHHs, in concordance with previous literature reports in other tissues [185, 186, 452, 453] and in hepatocytes [111]. Moreover, similar to the observations in lipid-loaded hepatocytes, aging affected the proportions of subpopulations of hepatocytes in the PHH *in vitro* culture, and aged hepatocytes showed increased oxidative stress and inflammation compared to young hepatocytes. In addition, aging and fat accumulation accelerated the dedifferentiation and expression loss process that PHHs undergo in monolayer *in vitro* culture. This indicates that intrinsic factors, as well as dietary factors, impact the functionality of the liver and the associated phenotype.

Taking the investigations further, the Sanofi-Aventis five-drug cocktail was administered to lipid-laden hepatocytes resembling the phenotype in NAFLD livers. In the four identified subgroups, fat accumulation decreased the capacity of single hepatocytes to metabolize drugs. When steatotic hepatocytes were exposed to a five-drug cocktail, regions upstream of *CYP450* enzyme involved in sterols metabolism and inflammatory processes were identified as differentially accessible, together

with solute carriers such as *SLC19A1* or *SLC7A2*. At the transcriptomic level, these genes were identified among the DEGs upon FFA+Cocktail vs. DMSO treatment at the transcriptomic level [375]. This indicates that exposure to a five-drug cocktail of cells containing fat accumulation leads to a rearrangement of the chromatin accessibility configuration of PHHs *in vitro*. Moreover, a significant downregulation of the main five CYPs was observed for subgroup III, which specialized in lipids and phase III metabolism compared to vehicle levels. This could play a major role in the development of adverse drug reactions and hepatotoxicity in especially susceptible population groups like patients with chronic liver diseases such as NAFLD, or the elderly. In the elderly population, coadministration of five or more drugs known as polypharmacy is a common clinical practice to treat age-related comorbidities. This highly increases the risk of the development of adverse drug reactions and drug-induced liver injury. Further studies on aging and disease, taking into consideration changes in the proportion of subpopulations of hepatocytes upon exposure to several drugs or fat accumulation simultaneously would help anticipate such unwanted events in early phases of drug development.

Taken together, the results presented show the plasticity and heterogeneity of the transcriptomic response of PHH *in vitro* toward intrinsic or extrinsic factors, which is regulated by the upstream chromatin configuration pattern under different environmental conditions.

6. Acknowledgements

I would like to begin by thanking my supervisor Dr. Celia Martínez Jiménez for giving me the opportunity to complete my doctoral thesis in her lab. For the fruitful and inspiring scientific discussions and problem-solving, for your guidance and useful both scientific and non-scientific advice, through struggles and doubtful times. I would like to extend my gratitude to the directorship of the Helmholtz Pioneer Campus: Prof. Dr. Maria Elena Torres Padilla, Prof. Dr. Vasilis Ntziachristos, Prof. Dr. Fabian Theis and Dr. Thomas Schwarz-Romond for the opportunity to join the institute. Special thanks to Thomas for believing in me and advocating joining the institute. Creating a multidisciplinary and collaborative environment is not an easy task.

I am grateful to Dr. Maria Colomé-Tatché for her great support, time, and guidance as a thesis committee mentor and in the scientific project we shared.

Special thanks to the Helmholtz Pioneer Campus team: Manuela Hartmann, Andreas Schröder, Stefanie Montag, and Almut Barden for enabling and facilitating projects to move forward, and for their managerial, regulatory, and administrative assistance.

Thanks to Dr. Cristina García-Caceres (Institute for Diabetes and Obesity) and her lab, especially to Dr. Raian Contreras, Dr. Tim Gruber, and Dr. Viktorian Miok for the prosperous and enjoyable collaboration in the oxytocin project.

I would also like to thank the Martínez-Jiménez lab members for the scientific exchange and the fun times. To Rizqah for sharing so much. To Patrick for his patient help on the scATAC-seq analysis journey; and Anna, for the assistance and laughter. I particularly want to thank Ioannis for his experimental support and for everything I learned from him about the single-cell world.

My deepest gratitude and appreciation to Maria Richter for the scientific and emotional support and for her selfless help and kindness. Working together was not only fruitful but also truly enjoyable, as were your fantastic baking skills. I enjoy our complementary sets of skills, personalities, and characters.

I wholeheartedly thank my incredible family and friends for their constant support and warmth, without whom none of this would have been possible. I am eternally grateful to you, Dani, for standing by me, for motivating me, and for your infinite patience, encouragement, and love.

7. References

1. Sanchez-Quant, E., et al., *Single-cell metabolic profiling reveals subgroups of primary human hepatocytes with heterogeneous responses to drug challenge.* Genome Biology, 2023. **24**(1): p. 234.
2. Aizarani, N., et al., *A human liver cell atlas reveals heterogeneity and epithelial progenitors.* Nature, 2019. **572**(7768): p. 199-204.
3. Liao, J., et al., *Single-cell RNA sequencing of human kidney.* Scientific Data, 2020. **7**(1): p. 4.
4. Litviňuková, M., et al., *Cells of the adult human heart.* Nature, 2020. **588**(7838): p. 466-472.
5. Halpern, K.B., et al., *Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells.* Nat Biotechnol, 2018. **36**(10): p. 962-970.
6. Halpern, K.B., et al., *Single-cell spatial reconstruction reveals global division of labour in the mammalian liver.* Nature, 2017. **542**(7641): p. 352-356.
7. Muraro, M.J., et al., *A Single-Cell Transcriptome Atlas of the Human Pancreas.* Cell Syst, 2016. **3**(4): p. 385-394.e3.
8. Ramachandran, P., et al., *Resolving the fibrotic niche of human liver cirrhosis at single-cell level.* Nature, 2019. **575**(7783): p. 512-518.
9. Reyfman, P.A., et al., *Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis.* Am J Respir Crit Care Med, 2019. **199**(12): p. 1517-1536.
10. Schaum, N., et al., *Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris.* Nature, 2018. **562**(7727): p. 367-372.
11. Regev, A., et al., *The Human Cell Atlas.* Elife, 2017. **6**.
12. Jones, R.C., et al., *The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans.* Science, 2022. **376**(6594): p. eabl4896.
13. The Tabula Sapiens, C., et al., *The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans.* Science. **376**(6594): p. eabl4896.
14. Wagner, A., A. Regev, and N. Yosef, *Revealing the vectors of cellular identity with single-cell genomics.* Nature biotechnology, 2016. **34**(11): p. 1145-1160.
15. Tanay, A. and A. Regev, *Scaling single-cell genomics from phenomenology to mechanism.* Nature, 2017. **541**(7637): p. 331-338.
16. Angelidis, I., et al., *An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics.* Nature Communications, 2019. **10**(1): p. 963.
17. Ben-Moshe, S., et al., *Spatial sorting enables comprehensive characterization of liver zonation.* Nature metabolism, 2019. **1**(9): p. 899-911.
18. Nadelmann, E.R., et al., *Isolation of Nuclei from Mammalian Cells and Tissues for Single-Nucleus Molecular Profiling.* Current Protocols, 2021. **1**(5): p. e132.
19. Richter, M.L., et al., *Single-nucleus RNA-seq2 reveals functional crosstalk between liver zonation and ploidy.* Nature Communications, 2021. **12**(1): p. 4264.
20. Rousselle, T.V., et al., *An optimized protocol for single nuclei isolation from clinical biopsies for RNA-seq.* Sci Rep, 2022. **12**(1): p. 9851.
21. Gross, A., et al., *Technologies for Single-Cell Isolation.* Int J Mol Sci, 2015. **16**(8): p. 16897-919.
22. Kolodziejczyk, Aleksandra A., et al., *The Technology and Biology of Single-Cell RNA Sequencing.* Molecular Cell, 2015. **58**(4): p. 610-620.
23. Hwang, B., J.H. Lee, and D. Bang, *Single-cell RNA sequencing technologies and bioinformaticspipelines.* Experimental & Molecular Medicine, 2018. **50**(8): p. 1-14.

24. Miltenyi, S., et al., *High gradient magnetic cell separation with MACS.* Cytometry, 1990. **11**(2): p. 231-238.
25. Habib, N., et al., *Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons.* Science, 2016. **353**(6302): p. 925-8.
26. Habib, N., et al., *Massively parallel single-nucleus RNA-seq with DroNc-seq.* Nat Methods, 2017. **14**(10): p. 955-958.
27. Habib, N., et al., *DroNc-Seq: Deciphering cell types in human archived brain tissues by massively-parallel single nucleus RNA-seq. bioRxiv: 115196.* 2017.
28. Haque, A., et al., *A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications.* Genome Medicine, 2017. **9**(1): p. 75.
29. Jovic, D., et al., *Single-cell RNA sequencing technologies and applications: A brief overview.* Clinical and Translational Medicine, 2022. **12**(3): p. e694.
30. Chen, G., B. Ning, and T. Shi, *Single-Cell RNA-Seq Technologies and Related Computational Data Analysis.* Front Genet, 2019. **10**: p. 317.
31. Picelli, S., et al., *Full-length RNA-seq from single cells using Smart-seq2.* Nature Protocols, 2014. **9**(1): p. 171-181.
32. Macosko, E.Z., et al., *Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets.* Cell, 2015. **161**(5): p. 1202-1214.
33. Gierahn, T.M., et al., *Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput.* Nature Methods, 2017. **14**(4): p. 395-398.
34. Zheng, G.X., et al., *Massively parallel digital transcriptional profiling of single cells.* Nat Commun, 2017. **8**: p. 14049.
35. Klein, Allon M., et al., *Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells.* Cell, 2015. **161**(5): p. 1187-1201.
36. Archer, N., et al., *Modeling Enzyme Processivity Reveals that RNA-Seq Libraries Are Biased in Characteristic and Correctable Ways.* Cell Syst, 2016. **3**(5): p. 467-479.e12.
37. Contreras, R.E., *Regulation of hypothalamic dopaminergic neurons and systemic eneregy metabolism by histone deacetylase 5.* Technische Universität München, 2021.
38. Baran-Gale, J., T. Chandra, and K. Kirschner, *Experimental design for single-cell RNA sequencing.* Briefings in Functional Genomics, 2018. **17**(4): p. 233-239.
39. Celton-Morizur, S. and C. Desdouets, *Polyploidization of liver cells.* Adv Exp Med Biol, 2010. **676**: p. 123-35.
40. Zhang, X., et al., *Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems.* Molecular Cell, 2019. **73**(1): p. 130-142.e5.
41. Islam, S., et al., *Quantitative single-cell RNA-seq with unique molecular identifiers.* Nature Methods, 2014. **11**(2): p. 163-166.
42. Katsuda, T., et al., *Transcriptomic Dissection of Hepatocyte Heterogeneity: Linking Ploidy, Zonation, and Stem/Progenitor Cell Characteristics.* Cellular and Molecular Gastroenterology and Hepatology, 2020. **9**(1): p. 161-183.
43. Kreutz, C., et al., *Hepatocyte Ploidy Is a Diversity Factor for Liver Homeostasis.* Front Physiol, 2017. **8**: p. 862.
44. Buenrostro, J.D., et al., *Single-cell chromatin accessibility reveals principles of regulatory variation.* Nature, 2015. **523**(7561): p. 486-90.
45. Cusanovich, D.A., et al., *A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility.* Cell, 2018. **174**(5): p. 1309-1324.e18.
46. Sinnamon, J.R., et al., *The accessible chromatin landscape of the murine hippocampus at single-cell resolution.* Genome Res, 2019. **29**(5): p. 857-869.
47. Satpathy, A.T., et al., *Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion.* Nature Biotechnology, 2019. **37**(8): p. 925-936.

48. Chen, T., et al., *Single-cell omics analysis reveals functional diversification of hepatocytes during liver regeneration.* JCI Insight, 2020. **5**(22).
49. Luo, C., et al., *Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex.* Science, 2017. **357**(6351): p. 600-604.
50. Li, Y.E., et al., *An atlas of gene regulatory elements in adult mouse cerebrum.* Nature, 2021. **598**(7879): p. 129-136.
51. Clark, S.J., et al., *Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity.* Genome Biology, 2016. **17**(1): p. 72.
52. Kamies, R. and C.P. Martinez-Jimenez, *Advances of single-cell genomics and epigenomics in human disease: where are we now?* Mamm Genome, 2020. **31**(5-6): p. 170-180.
53. Buenrostro, J.D., et al., *Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position.* Nature Methods, 2013. **10**(12): p. 1213-1218.
54. Stergachis, A.B., et al., *Developmental fate and cellular maturity encoded in human regulatory DNA landscapes.* Cell, 2013. **154**(4): p. 888-903.
55. Thurman, R.E., et al., *The accessible chromatin landscape of the human genome.* Nature, 2012. **489**(7414): p. 75-82.
56. Cusanovich, D.A., et al., *Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing.* Science, 2015. **348**(6237): p. 910-4.
57. Goryshin, I.Y. and W.S. Reznikoff, *Tn5 in vitro transposition.* J Biol Chem, 1998. **273**(13): p. 7367-74.
58. Adey, A., et al., *Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition.* Genome Biol, 2010. **11**(12): p. R119.
59. Clark, D.J., *Nucleosome positioning, nucleosome spacing and the nucleosome code.* J Biomol Struct Dyn, 2010. **27**(6): p. 781-93.
60. Chen, X., et al., *A rapid and robust method for single cell chromatin accessibility profiling.* Nature Communications, 2018. **9**(1): p. 5345.
61. Mezger, A., et al., *High-throughput chromatin accessibility profiling at single-cell resolution.* Nature Communications, 2018. **9**(1): p. 3647.
62. Lareau, C.A., et al., *Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility.* Nature Biotechnology, 2019. **37**(8): p. 916-924.
63. Dam, T.V., N.I. Toft, and L. Grøntved, *Cell-Type Resolved Insights into the Cis-Regulatory Genome of NAFLD.* Cells, 2022. **11**(5).
64. Krishnaswami, S.R., et al., *Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons.* Nat Protoc, 2016. **11**(3): p. 499-524.
65. Corces, M.R., et al., *An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues.* Nat Methods, 2017. **14**(10): p. 959-962.
66. Corces, M.R., et al., *Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution.* Nat Genet, 2016. **48**(10): p. 1193-203.
67. Bao, X., et al., *A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63.* Genome Biol, 2015. **16**: p. 284.
68. Genomics, X., *Demonstrated Protocol – Nuclei Isolation for Single Cell ATAC Sequencing • Rev D.* 2021.
69. Martínez-Jiménez, C.P., et al., *Underexpressed coactivators PGC1alpha and SRC1 impair hepatocyte nuclear factor 4 alpha function and promote dedifferentiation in human hepatoma cells.* J Biol Chem, 2006. **281**(40): p. 29840-9.
70. Martinez-Jimenez, et al., *Transcriptional Regulation and Expression of CYP3A4 in Hepatocytes.* Current Drug Metabolism, 2007. **8**(2): p. 185-194.

71. Martinez-Jimenez, C.P., et al., *Hepatocyte nuclear factor 4alpha coordinates a transcription factor network regulating hepatic fatty acid metabolism.* Molecular and cellular biology, 2010. **30**(3): p. 565-577.
72. Schulze, R.J., et al., *The cell biology of the hepatocyte: A membrane trafficking machine.* J Cell Biol, 2019. **218**(7): p. 2096-2112.
73. O'Hare, J.D. and A. Zsombok, *Brain-liver connections: role of the preautonomic PVN neurons.* American Journal of Physiology-Endocrinology and Metabolism, 2015. **310**(3): p. E183-E189.
74. Stanley, S., et al., *Identification of neuronal subpopulations that project from hypothalamus to both liver and adipose tissue polysynaptically.* Proc Natl Acad Sci U S A, 2010. **107**(15): p. 7024-9.
75. Garfield, A.S., et al., *A parabrachial-hypothalamic cholecystokinin neurocircuit controls counterregulatory responses to hypoglycemia.* Cell Metab, 2014. **20**(6): p. 1030-7.
76. Campbell, J.N., et al., *A molecular census of arcuate hypothalamus and median eminence cell types.* Nat Neurosci, 2017. **20**(3): p. 484-496.
77. Romanov, R.A., et al., *Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes.* Nat Neurosci, 2017. **20**(2): p. 176-188.
78. Lam, B.Y.H., et al., *Heterogeneity of hypothalamic pro-opiomelanocortin-expressing neurons revealed by single-cell RNA sequencing.* Mol Metab, 2017. **6**(5): p. 383-392.
79. Chen, R., et al., *Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity.* Cell Rep, 2017. **18**(13): p. 3227-3241.
80. Cowley, M.A., et al., *Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus.* Nature, 2001. **411**(6836): p. 480-4.
81. Jones, R.C., et al., *The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans.* Science, 2022. **376**(6594): p. eabl4896.
82. French, S.J., et al., *Adaptation to high-fat diets: effects on eating behaviour and plasma cholecystokinin.* Br J Nutr, 1995. **73**(2): p. 179-89.
83. Seo, W. and W.I. Jeong, *Hepatic non-parenchymal cells: Master regulators of alcoholic liver disease?* World J Gastroenterol, 2016. **22**(4): p. 1348-56.
84. MacParland, S.A., et al., *Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations.* Nature Communications, 2018. **9**(1): p. 4383.
85. McEnerney, L., et al., *Dual modulation of human hepatic zonation via canonical and non-canonical Wnt pathways.* Experimental & molecular medicine, 2017. **49**(12): p. e413-e413.
86. Martínez-Jiménez, C.P., et al., *Transcriptional regulation and expression of CYP3A4 in hepatocytes.* Curr Drug Metab, 2007. **8**(2): p. 185-94.
87. Garnier, D., et al., *Expansion of human primary hepatocytes in vitro through their amplification as liver progenitors in a 3D organoid system.* Scientific Reports, 2018. **8**(1): p. 8222.
88. Duncan, A.W., et al., *The ploidy conveyor of mature hepatocytes as a source of genetic variation.* Nature, 2010. **467**(7316): p. 707-710.
89. Duncan, A.W., et al., *Frequent aneuploidy among normal human hepatocytes.* Gastroenterology, 2012. **142**(1): p. 25-8.
90. Donne, R., et al., *Polyploidy in liver development, homeostasis and disease.* Nat Rev Gastroenterol Hepatol, 2020. **17**(7): p. 391-405.
91. Otto, S.P., *The evolutionary consequences of polyploidy.* Cell, 2007. **131**(3): p. 452-62.
92. Duncan, A.W., *Aneuploidy, polyploidy and ploidy reversal in the liver.* Semin Cell Dev Biol, 2013. **24**(4): p. 347-56.
93. Gentric, G. and C. Desdouets, *Polyploidization in liver tissue.* Am J Pathol, 2014. **184**(2): p. 322-31.
94. Bou-Nader, M., et al., *Polyploidy spectrum: a new marker in HCC classification.* Gut, 2020. **69**(2): p. 355-364.

95. Guidotti, J.E., et al., *Liver cell polyploidization: a pivotal role for binuclear hepatocytes.* J Biol Chem, 2003. **278**(21): p. 19095-101.
96. Kudryavtsev, B.N., et al., *Human hepatocyte polyploidization kinetics in the course of life cycle.* Virchows Arch B Cell Pathol Incl Mol Pathol, 1993. **64**(6): p. 387-93.
97. Toyoda, H., et al., *Conserved balance of hepatocyte nuclear DNA content in mononuclear and binuclear hepatocyte populations during the course of chronic viral hepatitis.* World J Gastroenterol, 2006. **12**(28): p. 4546-8.
98. Toyoda, H., et al., *Changes to hepatocyte ploidy and binuclearity profiles during human chronic viral hepatitis.* Gut, 2005. **54**(2): p. 297-302.
99. Cunningham, R.P. and N. Porat-Shliom, *Liver Zonation - Revisiting Old Questions With New Technologies.* Front Physiol, 2021. **12**: p. 732929.
100. Gebhardt, R. and M. Matz-Soja, *Liver zonation: Novel aspects of its regulation and its impact on homeostasis.* World J Gastroenterol, 2014. **20**(26): p. 8491-504.
101. Birchmeier, W., *Orchestrating Wnt signalling for metabolic liver zonation.* Nature Cell Biology, 2016. **18**(5): p. 463-465.
102. Wei, Y., et al., *Liver homeostasis is maintained by midlobular zone 2 hepatocytes.* Science, 2021. **371**(6532).
103. Kietzmann, T., *Metabolic zonation of the liver: The oxygen gradient revisited.* Redox biology, 2017. **11**: p. 622-630.
104. Manco, R. and S. Itzkovitz, *Liver zonation.* Journal of Hepatology, 2021. **74**(2): p. 466-468.
105. Ben-Moshe, S. and S. Itzkovitz, *Spatial heterogeneity in the mammalian liver.* Nature Reviews Gastroenterology & Hepatology, 2019. **16**(7): p. 395-410.
106. Park, S.R., et al., *Holistic characterization of single-hepatocyte transcriptome responses to high-fat diet.* American Journal of Physiology-Endocrinology and Metabolism, 2020. **320**(2): p. E244-E258.
107. Segal, J.M., et al., *Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors.* Nat Commun, 2019. **10**(1): p. 3350.
108. Tamburini, B.A.J., et al., *Chronic Liver Disease in Humans Causes Expansion and Differentiation of Liver Lymphatic Endothelial Cells.* Front Immunol, 2019. **10**: p. 1036.
109. Brosch, M., et al., *Epigenomic map of human liver reveals principles of zonated morphogenic and metabolic control.* Nature Communications, 2018. **9**(1): p. 4150.
110. Gravina, S., et al., *Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome.* Genome Biol, 2016. **17**(1): p. 150.
111. Chrysa, N., et al., *Single-cell resolution unravels spatial alterations in metabolism, transcriptome and epigenome of ageing liver.* bioRxiv, 2021: p. 2021.12.14.472593.
112. Patterson, A.D., F.J. Gonzalez, and J.R. Idle, *Xenobiotic metabolism: a view through the metabolometer.* Chem Res Toxicol, 2010. **23**(5): p. 851-60.
113. Johnson, C.H., et al., *Xenobiotic metabolomics: major impact on the metabolome.* Annu Rev Pharmacol Toxicol, 2012. **52**: p. 37-56.
114. Phang-Lyn, S. and V.A. Llerena, *Biochemistry, Biotransformation*, in *StatPearls*. 2022, StatPearls Publishing
Copyright © 2022, StatPearls Publishing LLC.: Treasure Island (FL).
115. Doogue, M.P. and T.M. Polasek, *The ABCD of clinical pharmacokinetics.* Ther Adv Drug Saf, 2013. **4**(1): p. 5-7.
116. Nelson, D.R., et al., *Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants.* Pharmacogenetics, 2004. **14**(1): p. 1-18.
117. Zanger, U.M. and M. Schwab, *Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation.* Pharmacology & Therapeutics, 2013. **138**(1): p. 103-141.

118. Guengerich, F.P., *Cytochrome p450 and chemical toxicology.* Chem Res Toxicol, 2008. **21**(1): p. 70-83.
119. Neve, E.P. and M. Ingelman-Sundberg, *Cytochrome P450 proteins: retention and distribution from the endoplasmic reticulum.* Curr Opin Drug Discov Devel, 2010. **13**(1): p. 78-85.
120. Wienkers, L.C. and T.G. Heath, *Predicting in vivo drug interactions from in vitro drug discovery data.* Nature Reviews Drug Discovery, 2005. **4**(10): p. 825-833.
121. Williams, J.A., et al., *Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUCi/AUC) ratios.* Drug Metab Dispos, 2004. **32**(11): p. 1201-8.
122. Spatzenegger, M. and W. Jaeger, *Clinical importance of hepatic cytochrome P450 in drug metabolism.* Drug Metab Rev, 1995. **27**(3): p. 397-417.
123. Rodriguez-Antona, C., et al., *Cytochrome P-450 mRNA expression in human liver and its relationship with enzyme activity.* Arch Biochem Biophys, 2001. **393**(2): p. 308-15.
124. Ingelman-Sundberg, M., *The human genome project and novel aspects of cytochrome P450 research.* Toxicol Appl Pharmacol, 2005. **207**(2 Suppl): p. 52-6.
125. (FDA), U.S.F.a.D.A. *Drug Development and Drug Interactions | Table of Substrates, Inhibitors and Inducers*. 2020 03/10/2020; Available from: https://www.fda.gov/drugs/drug-interactions-labeling/drug-development-and-drug-interactions-table-substrates-inhibitors-and-inducers.
126. Manikandan, P. and S. Nagini, *Cytochrome P450 Structure, Function and Clinical Significance: A Review.* Curr Drug Targets, 2018. **19**(1): p. 38-54.
127. Guengerich, F.P., M.R. Waterman, and M. Egli, *Recent Structural Insights into Cytochrome P450 Function.* Trends Pharmacol Sci, 2016. **37**(8): p. 625-640.
128. Kivistö, K.T. and H.K. Kroemer, *Use of probe drugs as predictors of drug metabolism in humans.* J Clin Pharmacol, 1997. **37**(S1): p. 40s-48s.
129. Kot, M. and W.A. Daniel, *Caffeine as a marker substrate for testing cytochrome P450 activity in human and rat.* Pharmacol Rep, 2008. **60**(6): p. 789-97.
130. Furuta, T., et al., *Influence of CYP2C19 pharmacogenetic polymorphism on proton pump inhibitor-based therapies.* Drug Metab Pharmacokinet, 2005. **20**(3): p. 153-67.
131. Linden, R., et al., *Relation between CYP2C19 phenotype and genotype in a group of Brazilian volunteers.* Brazilian Journal of Pharmaceutical Sciences, 2009. **45**: p. 461-467.
132. Andersson, T., et al., *Identification of human liver cytochrome P450 isoforms mediating omeprazole metabolism.* Br J Clin Pharmacol, 1993. **36**(6): p. 521-30.
133. Chiba, K., et al., *Oxidative metabolism of omeprazole in human liver microsomes: cosegregation with S-mephenytoin 4'-hydroxylation.* J Pharmacol Exp Ther, 1993. **266**(1): p. 52-9.
134. Rettie, A.E., et al., *Hydroxylation of warfarin by human cDNA-expressed cytochrome P-450: a role for P-4502C9 in the etiology of (S)-warfarin-drug interactions.* Chem Res Toxicol, 1992. **5**(1): p. 54-9.
135. Kim, S.Y., et al., *Metabolism of R- and S-warfarin by CYP2C19 into four hydroxywarfarins.* Drug Metab Lett, 2012. **6**(3): p. 157-64.
136. Kaminsky, L.S. and Z.Y. Zhang, *Human P450 metabolism of warfarin.* Pharmacol Ther, 1997. **73**(1): p. 67-74.
137. Berger, B., et al., *Cytochrome P450 Enzymes Involved in Metoprolol Metabolism and Use of Metoprolol as a CYP2D6 Phenotyping Probe Drug.* Frontiers in Pharmacology, 2018. **9**.
138. Brocker, C.N., et al., *Metabolomic profiling of metoprolol hypertension treatment reveals altered gut microbiota-derived urinary metabolites.* Hum Genomics, 2020. **14**(1): p. 10.
139. Bu, H.Z., *A literature review of enzyme kinetic parameters for CYP3A4-mediated metabolic reactions of 113 drugs in human liver microsomes: structure-kinetics relationship assessment.* Curr Drug Metab, 2006. **7**(3): p. 231-49.

140. Liu, Y.T., et al., *Drugs as CYP3A probes, inducers, and inhibitors.* Drug Metab Rev, 2007. **39**(4): p. 699-721.
141. Zanger, U.M., et al., *Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation.* Anal Bioanal Chem, 2008. **392**(6): p. 1093-108.
142. van Groen, B.D., et al., *The Oral Bioavailability and Metabolism of Midazolam in Stable Critically Ill Children: A Pharmacokinetic Microtracing Study.* Clinical Pharmacology & Therapeutics, 2021. **109**(1): p. 140-149.
143. Williams, J.A., et al., *Comparative metabolic capabilities of CYP3A4, CYP3A5, and CYP3A7.* Drug Metab Dispos, 2002. **30**(8): p. 883-91.
144. Janssen, A.W.F., et al., *Cytochrome P450 expression, induction and activity in human induced pluripotent stem cell-derived intestinal organoids and comparison with primary human intestinal epithelial cells and Caco-2 cells.* Archives of Toxicology, 2021. **95**(3): p. 907-922.
145. Zhou, S.F., J.P. Liu, and B. Chowbay, *Polymorphism of human cytochrome P450 enzymes and its clinical impact.* Drug Metab Rev, 2009. **41**(2): p. 89-295.
146. Lauschke, V.M. and M. Ingelman-Sundberg, *The Importance of Patient-Specific Factors for Hepatic Drug Response and Toxicity.* Int J Mol Sci, 2016. **17**(10).
147. Eichelbaum, M., M. Ingelman-Sundberg, and W.E. Evans, *Pharmacogenomics and individualized drug therapy.* Annu Rev Med, 2006. **57**: p. 119-37.
148. Preissner, S.C., et al., *Polymorphic cytochrome P450 enzymes (CYPs) and their role in personalized therapy.* PLoS One, 2013. **8**(12): p. e82562.
149. McGraw, J. and D. Waller, *Cytochrome P450 variations in different ethnic populations.* Expert Opin Drug Metab Toxicol, 2012. **8**(3): p. 371-82.
150. Myrand, S.P., et al., *Pharmacokinetics/genotype associations for major cytochrome P450 enzymes in native and first- and third-generation Japanese populations: comparison with Korean, Chinese, and Caucasian populations.* Clin Pharmacol Ther, 2008. **84**(3): p. 347-61.
151. Ingelman-Sundberg, M. and A. Gomez, *The past, present and future of pharmacoepigenomics.* Pharmacogenomics, 2010. **11**(5): p. 625-7.
152. Ingelman-Sundberg, M., et al., *Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects.* Pharmacol Ther, 2007. **116**(3): p. 496-526.
153. Dean, L. and M. Kane, *Omeprazole Therapy and CYP2C19 Genotype*, in *Medical Genetics Summaries*, V.M. Pratt, et al., Editors. 2012, National Center for Biotechnology Information (US): Bethesda (MD).
154. Mizutani, T., *PM frequencies of major CYPs in Asians and Caucasians.* Drug metabolism reviews, 2003. **35**(2-3): p. 99-106.
155. Johansson, I. and M. Ingelman-Sundberg, *Genetic polymorphism and toxicology--with emphasis on cytochrome p450.* Toxicol Sci, 2011. **120**(1): p. 1-13.
156. Jancova, P., P. Anzenbacher, and E. Anzenbacherova, *Phase II drug metabolizing enzymes.* Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2010. **154**(2): p. 103-16.
157. Jancova, P., P. Anzenbacher, and E. Anzenbacherova, *PHASE II DRUG METABOLIZING ENZYMES.* Biomedical papers, 2010. **154**(2): p. 103-116.
158. Petra, J. and Š. Michal, *Phase II Drug Metabolism*, in *Topics on Drug Metabolism*, P. James, Editor. 2012, IntechOpen: Rijeka. p. Ch. 2.
159. Almazroo, O.A., M.K. Miah, and R. Venkataramanan, *Drug Metabolism in the Liver.* Clin Liver Dis, 2017. **21**(1): p. 1-20.
160. Esteves, F., J. Rueff, and M. Kranendonk, *The Central Role of Cytochrome P450 in Xenobiotic Metabolism—A Brief Review on a Fascinating Enzyme Family.* Journal of Xenobiotics, 2021. **11**(3).

161. Omiecinski, C.J., et al., *Xenobiotic Metabolism, Disposition, and Regulation by Receptors: From Biochemical Phenomenon to Predictors of Major Toxicities.* Toxicological Sciences, 2011. **120**(suppl_1): p. S49-S75.
162. Shao, H., L.Q. Chen, and J. Xu, *Treatment of dyslipidemia in the elderly.* J Geriatr Cardiol, 2011. **8**(1): p. 55-64.
163. Chang, T.I., et al., *Polypharmacy, hospitalization, and mortality risk: a nationwide cohort study.* Scientific Reports, 2020. **10**(1): p. 18964.
164. Munday, D., et al., *The prevalence of non-communicable disease in older people in prison: a systematic review and meta-analysis.* Age and Ageing, 2019. **48**(2): p. 204-212.
165. Johnell, K. and I. Klarin, *The relationship between number of drugs and potential drug-drug interactions in the elderly: a study of over 600,000 elderly patients from the Swedish Prescribed Drug Register.* Drug Saf, 2007. **30**(10): p. 911-8.
166. *American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults.* J Am Geriatr Soc, 2012. **60**(4): p. 616-31.
167. Beijer, H.J. and C.J. de Blaey, *Hospitalisations caused by adverse drug reactions (ADR): a meta-analysis of observational studies.* Pharm World Sci, 2002. **24**(2): p. 46-54.
168. Mehta, R.S., et al., *Emerging approaches to polypharmacy among older adults.* Nature Aging, 2021. **1**(4): p. 347-356.
169. Lopez-Otin, C., et al., *The hallmarks of aging.* Cell, 2013. **153**(6): p. 1194-217.
170. Hunt, N.J., et al., *Hallmarks of Aging in the Liver.* Comput Struct Biotechnol J, 2019. **17**: p. 1151-1161.
171. Alberti, K.G. and P.Z. Zimmet, *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation.* Diabet Med, 1998. **15**(7): p. 539-53.
172. Saklayen, M.G., *The Global Epidemic of the Metabolic Syndrome.* Curr Hypertens Rep, 2018. **20**(2): p. 12.
173. Ogrodnik, M., et al., *Cellular senescence drives age-dependent hepatic steatosis.* Nature Communications, 2017. **8**(1): p. 15691.
174. Baiocchi, L., et al., *Impact of Aging on Liver Cells and Liver Disease: Focus on the Biliary and Vascular Compartments.* Hepatology Communications, 2021. **5**(7): p. 1125-1137.
175. Schmucker, D.L., *Age-related changes in liver structure and function: Implications for disease ?* Exp Gerontol, 2005. **40**(8-9): p. 650-9.
176. Duval, C., et al., *Adipose Tissue Dysfunction Signals Progression of Hepatic Steatosis Towards Nonalcoholic Steatohepatitis in C57Bl/6 Mice.* Diabetes, 2010. **59**(12): p. 3181-3191.
177. Sheedfar, F., et al., *Liver diseases and aging: friends or foes?* Aging Cell, 2013. **12**(6): p. 950-4.
178. Gong, Z., et al., *Hepatic lipid metabolism and non-alcoholic fatty liver disease in aging.* Molecular and Cellular Endocrinology, 2017. **455**: p. 115-130.
179. Golabi, P., et al., *Prevalence and long-term outcomes of non-alcoholic fatty liver disease among elderly individuals from the United States.* BMC Gastroenterology, 2019. **19**(1): p. 56.
180. Chung, K.W., *Advances in Understanding of the Role of Lipid Metabolism in Aging.* Cells, 2021. **10**(4).
181. González-Rodríguez, Á., et al., *Impaired autophagic flux is associated with increased endoplasmic reticulum stress during the development of NAFLD.* Cell Death & Disease, 2014. **5**(4): p. e1179-e1179.
182. Zhang, L. and H.-H. Wang, *The essential functions of endoplasmic reticulum chaperones in hepatic lipid metabolism.* Digestive and Liver Disease, 2016. **48**(7): p. 709-716.
183. Kumar, A., et al., *Patients with Nonalcoholic Fatty Liver Disease (NAFLD) have Higher Oxidative Stress in Comparison to Chronic Viral Hepatitis.* Journal of Clinical and Experimental Hepatology, 2013. **3**(1): p. 12-18.

184. Nikopoulou, C., S. Parekh, and P. Tessarz, *Ageing and sources of transcriptional heterogeneity.* Biological Chemistry, 2019. **400**(7): p. 867-878.
185. Martinez-Jimenez, C.P., et al., *Aging increases cell-to-cell transcriptional variability upon immune stimulation.* Science, 2017. **355**(6332): p. 1433-1436.
186. Bahar, R., et al., *Increased cell-to-cell variation in gene expression in ageing mouse heart.* Nature, 2006. **441**(7096): p. 1011-4.
187. Enge, M., et al., *Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns.* Cell, 2017. **171**(2): p. 321-330.e14.
188. Almanzar, N., et al., *A single-cell transcriptomic atlas characterizes ageing tissues in the mouse.* Nature, 2020. **583**(7817): p. 590-595.
189. Covarrubias, A.J., et al., *Senescent cells promote tissue NAD+ decline during ageing via the activation of CD38+ macrophages.* Nature Metabolism, 2020. **2**(11): p. 1265-1283.
190. Alqahtani, S.A. and J.M. Schattenberg, *NAFLD in the Elderly.* Clin Interv Aging, 2021. **16**: p. 1633-1649.
191. Bozukova, M., et al., *Aging is associated with increased chromatin accessibility and reduced polymerase pausing in liver.* Mol Syst Biol, 2022. **18**(9): p. e11002.
192. Cheung, P., et al., *Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging.* Cell, 2018. **173**(6): p. 1385-1397.e14.
193. Nikopoulou, C., et al., *Single-cell resolution unravels spatial alterations in metabolism, transcriptome and epigenome of ageing liver.* bioRxiv, 2021: p. 2021.12.14.472593.
194. Chalasani, N., et al., *The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association.* Hepatology, 2012. **55**(6): p. 2005-23.
195. Chalasani, N., et al., *Features and Outcomes of 899 Patients With Drug-Induced Liver Injury: The DILIN Prospective Study.* Gastroenterology, 2015. **148**(7): p. 1340-52.e7.
196. Leoni, S., et al., *Current guidelines for the management of non-alcoholic fatty liver disease: A systematic review with comparative analysis.* World J Gastroenterol, 2018. **24**(30): p. 3361-3373.
197. Younossi, Z.M., et al., *Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.* Hepatology, 2016. **64**(1): p. 73-84.
198. Begriche, K., et al., *Drug-induced toxicity on mitochondria and lipid metabolism: mechanistic diversity and deleterious consequences for the liver.* J Hepatol, 2011. **54**(4): p. 773-94.
199. Marchesini, G., et al., *Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome.* Hepatology, 2003. **37**(4): p. 917-23.
200. Machado, M., P. Marques-Vidal, and H. Cortez-Pinto, *Hepatic histology in obese patients undergoing bariatric surgery.* J Hepatol, 2006. **45**(4): p. 600-6.
201. Adams, L.A., et al., *The natural history of nonalcoholic fatty liver disease: a population-based cohort study.* Gastroenterology, 2005. **129**(1): p. 113-21.
202. Geisler, C.E. and B.J. Renquist, *Hepatic lipid accumulation: cause and consequence of dysregulated glucoregulatory hormones.* J Endocrinol, 2017. **234**(1): p. R1-r21.
203. Hliwa, A., et al., *The Role of Fatty Acids in Non-Alcoholic Fatty Liver Disease Progression: An Update.* Int J Mol Sci, 2021. **22**(13).
204. Perazzo, H. and J.-F. Dufour, *The therapeutic landscape of non-alcoholic steatohepatitis.* Liver International, 2017. **37**(5): p. 634-647.
205. Zhang, X.Q., et al., *Role of endoplasmic reticulum stress in the pathogenesis of nonalcoholic fatty liver disease.* World J Gastroenterol, 2014. **20**(7): p. 1768-76.
206. Delli Bovi, A.P., et al., *Oxidative Stress in Non-alcoholic Fatty Liver Disease. An Updated Mini Review.* Front Med (Lausanne), 2021. **8**: p. 595371.

207. Fisher, C.D., et al., *Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of nonalcoholic fatty liver disease.* Drug Metab Dispos, 2009. **37**(10): p. 2087-94.
208. Abdelmegeed, M.A., et al., *Cytochrome P450-2E1 promotes aging-related hepatic steatosis, apoptosis and fibrosis through increased nitroxidative stress.* Free Radic Biol Med, 2016. **91**: p. 188-202.
209. Aubert, J., et al., *Increased expression of cytochrome P450 2E1 in nonalcoholic fatty liver disease: mechanisms and pathophysiological role.* Clin Res Hepatol Gastroenterol, 2011. **35**(10): p. 630-7.
210. Su, Q., et al., *Single-cell RNA transcriptome landscape of hepatocytes and non-parenchymal cells in healthy and NAFLD mouse liver.* iScience, 2021. **24**(11): p. 103233.
211. Xiong, X., et al., *Landscape of Intercellular Crosstalk in Healthy and NASH Liver Revealed by Single-Cell Secretome Gene Analysis.* Molecular Cell, 2019. **75**(3): p. 644-660.e5.
212. Fred, R.G., et al., *Single-cell transcriptome and cell type-specific molecular pathways of human non-alcoholic steatohepatitis.* Scientific Reports, 2022. **12**(1): p. 13484.
213. Wang, Z.-Y., et al., *Single-cell and bulk transcriptomics of the liver reveals potential targets of NASH with fibrosis.* Scientific Reports, 2021. **11**(1): p. 19396.
214. Rodríguez-Sanabria, J.S., et al., *An Update in Epigenetics in Metabolic-Associated Fatty Liver Disease.* Frontiers in Medicine, 2022. **8**.
215. Takeuchi, F., et al., *Gene-regulation modules in nonalcoholic fatty liver disease revealed by single-nucleus ATAC-seq.* Life Sci Alliance, 2023. **6**(10).
216. Fumihiko, T., et al., *Single-nucleus ATAC-seq elucidates major modules of gene regulation in the development of non-alcoholic fatty liver disease.* bioRxiv, 2022: p. 2022.07.12.499681.
217. Kang, B., et al., *The Chromatin Accessibility Landscape of Nonalcoholic Fatty Liver Disease Progression.* Mol Cells, 2022. **45**(5): p. 343-352.
218. Seebacher, F., et al., *Hepatic lipid droplet homeostasis and fatty liver disease.* Semin Cell Dev Biol, 2020. **108**: p. 72-81.
219. Friedman, S.L., et al., *Mechanisms of NAFLD development and therapeutic strategies.* Nature Medicine, 2018. **24**(7): p. 908-922.
220. Merrell, M.D. and N.J. Cherrington, *Drug metabolism alterations in nonalcoholic fatty liver disease.* Drug metabolism reviews, 2011. **43**(3): p. 317-334.
221. Massart, J., et al., *Role of nonalcoholic fatty liver disease as risk factor for drug-induced hepatotoxicity.* J Clin Transl Res, 2017. **3**(Suppl 1): p. 212-232.
222. Kostrzewski T, C.T., Snow SA, Ouro-Gnao L, Rowe C, Large EM, Hughes DJ., *Three-dimensional perfused human in vitro model of non-alcoholic fatty liver disease.* World J Gastroenterol, 2017. **23(2)**: p. 204-215.
223. Cobbina, E. and F. Akhlaghi, *Non-alcoholic fatty liver disease (NAFLD) - pathogenesis, classification, and effect on drug metabolizing enzymes and transporters.* Drug Metab Rev, 2017. **49**(2): p. 197-211.
224. Jamwal, R., et al., *Nonalcoholic Fatty Liver Disease and Diabetes Are Associated with Decreased CYP3A4 Protein Expression and Activity in Human Liver.* Molecular Pharmaceutics, 2018. **15**(7): p. 2621-2632.
225. Donato, M.T., et al., *Potential impact of steatosis on cytochrome P450 enzymes of human hepatocytes isolated from fatty liver grafts.* Drug Metab Dispos, 2006. **34**(9): p. 1556-62.
226. Merrell, M.D. and N.J. Cherrington, *Drug metabolism alterations in nonalcoholic fatty liver disease.* Drug Metab Rev, 2011. **43**(3): p. 317-34.
227. Younossi, Z.M., et al., *A genomic and proteomic study of the spectrum of nonalcoholic fatty liver disease.* Hepatology, 2005. **42**(3): p. 665-74.
228. Rey-Bedon, C., et al., *CYP450 drug inducibility in NAFLD via an in vitro hepatic model: Understanding drug-drug interactions in the fatty liver.* Biomedicine & Pharmacotherapy, 2022. **146**: p. 112377.

229. Gomez-Lechon, M.J., J.V. Castell, and M.T. Donato, *An update on metabolism studies using human hepatocytes in primary culture.* Expert Opin Drug Metab Toxicol, 2008. **4**(7): p. 837-54.
230. Donato, M.T., et al., *Cell lines: a tool for in vitro drug metabolism studies.* Curr Drug Metab, 2008. **9**(1): p. 1-11.
231. Tascher, G., et al., *In-Depth Proteome Analysis Highlights HepaRG Cells as a Versatile Cell System Surrogate for Primary Human Hepatocytes.* Cells, 2019. **8**(2).
232. Donato, M.T., L. Tolosa, and M.J. Gómez-Lechón, *Culture and Functional Characterization of Human Hepatoma HepG2 Cells.* Methods Mol Biol, 2015. **1250**: p. 77-93.
233. Andersson, T.B., K.P. Kanebratt, and J.G. Kenna, *The HepaRG cell line: a unique in vitro tool for understanding drug metabolism and toxicology in human.* Expert Opin Drug Metab Toxicol, 2012. **8**(7): p. 909-20.
234. Kanno, Y., et al., *Differences in Gene Regulation by Dual Ligands of Nuclear Receptors Constitutive Androstane Receptor (CAR) and Pregnane X Receptor (PXR) in HepG2 Cells Stably Expressing CAR/PXR.* Drug Metab Dispos, 2016. **44**(8): p. 1158-63.
235. Hernandez, J.P., L.C. Mota, and W.S. Baldwin, *Activation of CAR and PXR by Dietary, Environmental and Occupational Chemicals Alters Drug Metabolism, Intermediary Metabolism, and Cell Proliferation.* Curr Pharmacogenomics Person Med, 2009. **7**(2): p. 81-105.
236. Tolosa, L., et al., *Advantageous use of HepaRG cells for the screening and mechanistic study of drug-induced steatosis.* Toxicology and Applied Pharmacology, 2016. **302**: p. 1-9.
237. Serras, A.S., et al., *A Critical Perspective on 3D Liver Models for Drug Metabolism and Toxicology Studies.* Frontiers in Cell and Developmental Biology, 2021. **9**(203).
238. Gómez-Lechón, M.J., et al., *Competency of different cell models to predict human hepatotoxic drugs.* Expert Opinion on Drug Metabolism & Toxicology, 2014. **10**(11): p. 1553-1568.
239. Soldatow, V.Y., et al., *In vitro models for liver toxicity testing.* Toxicology research, 2013. **2**(1): p. 23-39.
240. Guillouzo, A., et al., *The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics.* Chem Biol Interact, 2007. **168**(1): p. 66-73.
241. Lübberstedt, M., et al., *HepaRG human hepatic cell line utility as a surrogate for primary human hepatocytes in drug metabolism assessment in vitro.* J Pharmacol Toxicol Methods, 2011. **63**(1): p. 59-68.
242. Gerets, H.H., et al., *Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins.* Cell Biol Toxicol, 2012. **28**(2): p. 69-87.
243. Gómez-Lechón, M.J., L. Tolosa, and M.T. Donato, *Upgrading HepG2 cells with adenoviral vectors that encode drug-metabolizing enzymes: application for drug hepatotoxicity testing.* Expert Opinion on Drug Metabolism & Toxicology, 2017. **13**(2): p. 137-148.
244. Eynaudi, A., et al., *Differential Effects of Oleic and Palmitic Acids on Lipid Droplet-Mitochondria Interaction in the Hepatic Cell Line HepG2.* Frontiers in Nutrition, 2021. **8**.
245. Kanebratt, K.P. and T.B. Andersson, *HepaRG cells as an in vitro model for evaluation of cytochrome P450 induction in humans.* Drug Metab Dispos, 2008. **36**(1): p. 137-45.
246. Kanebratt, K.P. and T.B. Andersson, *Evaluation of HepaRG cells as an in vitro model for human drug metabolism studies.* Drug Metab Dispos, 2008. **36**(7): p. 1444-52.
247. Jossé, R., et al., *Long-term functional stability of human HepaRG hepatocytes and use for chronic toxicity and genotoxicity studies.* Drug Metab Dispos, 2008. **36**(6): p. 1111-8.
248. Le Vee, M., et al., *Functional expression of sinusoidal and canalicular hepatic drug transporters in the differentiated human hepatoma HepaRG cell line.* European Journal of Pharmaceutical Sciences, 2006. **28**(1): p. 109-117.

249. Seglen, P.O., *Preparation of rat liver cells: III. Enzymatic requirements for tissue dispersion.* Experimental Cell Research, 1973. **82**(2): p. 391-398.
250. Guguen-Guillouzo, C., et al., *High yield preparation of isolated human adult hepatocytes by enzymatic perfusion of the liver.* Cell Biol Int Rep, 1982. **6**(6): p. 625-8.
251. Godoy, P., 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.* Arch Toxicol, 2013. **87**(8): p. 1315-530.
252. Zeilinger, K., et al., *Cell sources for in vitro human liver cell culture models.* Exp Biol Med (Maywood), 2016. **241**(15): p. 1684-98.
253. Heslop, J.A., et al., *Mechanistic evaluation of primary human hepatocyte culture using global proteomic analysis reveals a selective dedifferentiation profile.* Arch Toxicol, 2017. **91**(1): p. 439-452.
254. Lauschke, V.M., et al., *Massive rearrangements of cellular MicroRNA signatures are key drivers of hepatocyte dedifferentiation.* Hepatology, 2016. **64**(5): p. 1743-1756.
255. Sison-Young, R.L., et al., *A multicenter assessment of single-cell models aligned to standard measures of cell health for prediction of acute hepatotoxicity.* Arch Toxicol, 2017. **91**(3): p. 1385-1400.
256. Scheidecker, B., et al., *Induction of in vitro Metabolic Zonation in Primary Hepatocytes Requires Both Near-Physiological Oxygen Concentration and Flux.* Front Bioeng Biotechnol, 2020. **8**: p. 524.
257. Rodríguez-Antona, C., et al., *Cytochrome P450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells.* Xenobiotica, 2002. **32**(6): p. 505-20.
258. Ohtsuki, S., et al., *Simultaneous absolute protein quantification of transporters, cytochromes P450, and UDP-glucuronosyltransferases as a novel approach for the characterization of individual human liver: comparison with mRNA levels and activities.* Drug metabolism and Disposition, 2012. **40**(1): p. 83-92.
259. Ben-Moshe, S. and S. Itzkovitz, *Spatial heterogeneity in the mammalian liver.* Nat Rev Gastroenterol Hepatol, 2019. **16**(7): p. 395-410.
260. Gómez-Lechón, M.J., et al., *A human hepatocellular in vitro model to investigate steatosis.* Chem Biol Interact, 2007. **165**(2): p. 106-16.
261. Rogue, A., et al., *PPAR agonists reduce steatosis in oleic acid-overloaded HepaRG cells.* Toxicol Appl Pharmacol, 2014. **276**(1): p. 73-81.
262. Kozyra, M., et al., *Human hepatic 3D spheroids as a model for steatosis and insulin resistance.* Sci Rep, 2018. **8**(1): p. 14297.
263. Wobser, H., et al., *Lipid accumulation in hepatocytes induces fibrogenic activation of hepatic stellate cells.* Cell Research, 2009. **19**(8): p. 996-1005.
264. Soret, P.A., et al., *In Vitro and In Vivo Models of Non-Alcoholic Fatty Liver Disease: A Critical Appraisal.* J Clin Med, 2020. **10**(1).
265. Aoudjehane, L., et al., *Novel defatting strategies reduce lipid accumulation in primary human culture models of liver steatosis.* Dis Model Mech, 2020. **13**(4).
266. Levy, G., M. Cohen, and Y. Nahmias, *In Vitro Cell Culture Models of Hepatic Steatosis.* Methods Mol Biol, 2015. **1250**: p. 377-90.
267. Mashek, D.G., *Hepatic lipid droplets: A balancing act between energy storage and metabolic dysfunction in NAFLD.* Molecular Metabolism, 2021. **50**: p. 101115.
268. Selen, E.S., J. Choi, and M.J. Wolfgang, *Discordant hepatic fatty acid oxidation and triglyceride hydrolysis leads to liver disease.* JCI Insight, 2021. **6**(2).
269. Alsabeeh, N., et al., *Cell culture models of fatty acid overload: Problems and solutions.* Biochim Biophys Acta Mol Cell Biol Lipids, 2018. **1863**(2): p. 143-151.

270. Müller, F.A. and S.J. Sturla, *Human in vitro models of nonalcoholic fatty liver disease.* Current Opinion in Toxicology, 2019. **16**: p. 9-16.
271. Castell, J.V., et al., *Hepatocyte cell lines: their use, scope and limitations in drug metabolism studies.* Expert Opin Drug Metab Toxicol, 2006. **2**(2): p. 183-212.
272. Fuhr, U., A. Jetter, and J. Kirchheiner, *Appropriate phenotyping procedures for drug metabolizing enzymes and transporters in humans and their simultaneous use in the "cocktail" approach.* Clin Pharmacol Ther, 2007. **81**(2): p. 270-83.
273. Snyder, B.D., et al., *Evaluation of felodipine as a potential perpetrator of pharmacokinetic drug-drug interactions.* European journal of clinical pharmacology, 2014. **70**(9): p. 1115-1122.
274. Spaggiari, D., et al., *Phenotyping of CYP450 in human liver microsomes using the cocktail approach.* Anal Bioanal Chem, 2014. **406**(20): p. 4875-87.
275. Chen, Z.-h., et al., *An improved substrate cocktail for assessing direct inhibition and time-dependent inhibition of multiple cytochrome P450s.* Acta Pharmacologica Sinica, 2016. **37**(5): p. 708-718.
276. Spaggiari, D., Y. Daali, and S. Rudaz, *An extensive cocktail approach for rapid risk assessment of in vitro CYP450 direct reversible inhibition by xenobiotic exposure.* Toxicology and Applied Pharmacology, 2016. **302**: p. 41-51.
277. Carrão, D.B., A.R.M. de Oliveira, and I.R.S. Magalhães, *Challenges of probe cocktail approach for human drug–drug interaction assays.* Bioanalysis, 2018. **10**(24): p. 1969-1972.
278. Hakkola, J., et al., *Inhibition and induction of CYP enzymes in humans: an update.* Archives of Toxicology, 2020. **94**(11): p. 3671-3722.
279. Streetman, D.S., et al., *Combined phenotypic assessment of CYP1A2, CYP2C19, CYP2D6, CYP3A, N-acetyltransferase-2, and xanthine oxidase with the "Cooperstown cocktail".* Clinical Pharmacology & Therapeutics, 2000. **68**(4): p. 375-383.
280. Streetman, D.S., J.S. Bertino Jr, and A.N. Nafziger, *Phenotyping of drug-metabolizing enzymes in adults: a review of in-vivo cytochrome P450 phenotyping probes.* Pharmacogenetics and Genomics, 2000. **10**(3): p. 187-216.
281. Christensen, M., et al., *The Karolinska cocktail for phenotyping of five human cytochrome P450 enzymes.* Clin Pharmacol Ther, 2003. **73**(6): p. 517-28.
282. Ryu, J.Y., et al., *Development of the "Inje cocktail" for high-throughput evaluation of five human cytochrome P450 isoforms in vivo.* Clin Pharmacol Ther, 2007. **82**(5): p. 531-40.
283. Bosilkovska, M., et al., *Geneva cocktail for cytochrome p450 and P-glycoprotein activity assessment using dried blood spots.* Clin Pharmacol Ther, 2014. **96**(3): p. 349-59.
284. Berger, B., et al., *Comparison of Liver Cell Models Using the Basel Phenotyping Cocktail.* Front Pharmacol, 2016. **7**: p. 443.
285. Turpault, S., et al., *Pharmacokinetic assessment of a five-probe cocktail for CYPs 1A2, 2C9, 2C19, 2D6 and 3A.* Br J Clin Pharmacol, 2009. **68**(6): p. 928-35.
286. (EMA), E.M.A., *Guideline on the investigation of drug interactions - Revision 1.*
287. de Vries, E.M., et al., *Fasting-Induced Changes in Hepatic P450 Mediated Drug Metabolism Are Largely Independent of the Constitutive Androstane Receptor CAR.* PLOS ONE, 2016. **11**(7): p. e0159552.
288. Hakooz, N.M., *Caffeine metabolic ratios for the in vivo evaluation of CYP1A2, N-acetyltransferase 2, xanthine oxidase and CYP2A6 enzymatic activities.* Curr Drug Metab, 2009. **10**(4): p. 329-38.
289. Astrid, N., *Interindividual Differences in Caffeine Metabolism and Factors Driving Caffeine Consumption.* Pharmacological Reviews, 2018. **70**(2): p. 384.
290. Gu, L., et al., *Biotransformation of caffeine, paraxanthine, theobromine and theophylline by cDNA-expressed human CYP1A2 and CYP2E1.* Pharmacogenetics, 1992. **2**(2): p. 73-7.
291. Tassaneeyakul, W., et al., *Caffeine metabolism by human hepatic cytochromes P450: contributions of 1A2, 2E1 and 3A isoforms.* Biochem Pharmacol, 1994. **47**(10): p. 1767-76.

292. Cameron, M.D., et al., *Cooperative binding of acetaminophen and caffeine within the P450 3A4 active site.* Chem Res Toxicol, 2007. **20**(10): p. 1434-41.
293. Cederberg, C., T. Andersson, and I. Skånberg, *Omeprazole: Pharmacokinetics and Metabolism in Man.* Scandinavian Journal of Gastroenterology, 1989. **24**(sup166): p. 33-40.
294. Shirasaka, Y., et al., *Inhibition of CYP2C19 and CYP3A4 by omeprazole metabolites and their contribution to drug-drug interactions.* Drug Metab Dispos, 2013. **41**(7): p. 1414-24.
295. Han, X.M., et al., *Inducibility of CYP1A2 by omeprazole in vivo related to the genetic polymorphism of CYP1A2.* Br J Clin Pharmacol, 2002. **54**(5): p. 540-3.
296. Takahashi, H., et al., *Comparisons between in-vitro and in-vivo metabolism of (S)-warfarin: catalytic activities of cDNA-expressed CYP2C9, its Leu359 variant and their mixture versus unbound clearance in patients with the corresponding CYP2C9 genotypes.* Pharmacogenetics, 1998. **8**(5): p. 365-73.
297. Kaminsky, L.S., et al., *Correlation of human cytochrome P4502C substrate specificities with primary structure: warfarin as a probe.* Mol Pharmacol, 1993. **43**(2): p. 234-9.
298. Jones, D.R., et al., *Contribution of three CYP3A isoforms to metabolism of R- and S-warfarin.* Drug Metab Lett, 2010. **4**(4): p. 213-9.
299. Lennard, M.S., et al., *Oxidation phenotype--a major determinant of metoprolol metabolism and response.* N Engl J Med, 1982. **307**(25): p. 1558-60.
300. Berger, B., et al., *Cytochrome P450 Enzymes Involved in Metoprolol Metabolism and Use of Metoprolol as a CYP2D6 Phenotyping Probe Drug.* Front Pharmacol, 2018. **9**: p. 774.
301. Borkar, R.M., et al., *An evaluation of the CYP2D6 and CYP3A4 inhibition potential of metoprolol metabolites and their contribution to drug-drug and drug-herb interaction by LC-ESI/MS/MS.* Biomed Chromatogr, 2016. **30**(10): p. 1556-72.
302. Kronbach, T., et al., *Oxidation of midazolam and triazolam by human liver cytochrome P450IIIA4.* Mol Pharmacol, 1989. **36**(1): p. 89-96.
303. van Waterschoot, R.A., et al., *Midazolam metabolism in cytochrome P450 3A knockout mice can be attributed to up-regulated CYP2C enzymes.* Mol Pharmacol, 2008. **73**(3): p. 1029-36.
304. Shida, S. and H. Yamazaki, *Human plasma concentrations of five cytochrome P450 probes extrapolated from pharmacokinetics in dogs and minipigs using physiologically based pharmacokinetic modeling.* Xenobiotica, 2016. **46**(9): p. 759-64.
305. Koyanagi, T., et al., *Age-related changes of hepatic clearances of cytochrome P450 probes, midazolam and R-/S-warfarin in combination with caffeine, omeprazole and metoprolol in cynomolgus monkeys using in vitro–in vivo correlation.* Xenobiotica, 2015. **45**(4): p. 312-321.
306. Gruber, T., et al., *High-calorie diets uncouple hypothalamic oxytocin neurons from a gut-to-brain satiation pathway via κ-opioid signaling.* Cell Rep, 2023. **42**(10): p. 113305.
307. Deal, R.B. and S. Henikoff, *The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana.* Nat Protoc, 2011. **6**(1): p. 56-68.
308. Rodrigues, O.R. and S. Monard, *A rapid method to verify single-cell deposition setup for cell sorters.* Cytometry A, 2016. **89**(6): p. 594-600.
309. Mora-Castilla, S., et al., *Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing.* SLAS Technology, 2016. **21**(4): p. 557-567.
310. Franzén, O., L.M. Gan, and J.L.M. Björkegren, *PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data.* Database (Oxford), 2019. **2019**.
311. Ramakers, C., et al., *Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.* Neurosci Lett, 2003. **339**(1): p. 62-6.
312. Bioscience, L. 2023; Available from: https://bioscience.lonza.com/lonza_bs/CH/en/ADME-and-Toxicology/p/00000000000222111/Human-Hepatocytes%2C-Cryopreserved%2C-Plateable-and-Interaction-Qualified.
313. Wolock, S.L., R. Lopez, and A.M. Klein, *Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data.* Cell Systems, 2019. **8**(4): p. 281-291.e9.

314. Canchola, J., *Correct Use of Percent Coefficient of Variation (%CV) Formula for Log-Transformed Data.* MedCrave Online Journal of Proteomics & Bioinformatics, 2017. **6**: p. 1-3.
315. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
316. Gaedigk, A., et al., *The Pharmacogene Variation (PharmVar) Consortium: Incorporation of the Human Cytochrome P450 (CYP) Allele Nomenclature Database.* Clin Pharmacol Ther, 2018. **103**(3): p. 399-401.
317. Landrum, M.J., et al., *ClinVar: public archive of relationships among sequence variation and human phenotype.* Nucleic Acids Res, 2014. **42**(Database issue): p. D980-5.
318. Landrum, M.J., et al., *ClinVar: public archive of interpretations of clinically relevant variants.* Nucleic Acids Res, 2016. **44**(D1): p. D862-8.
319. Sherry, S.T., et al., *dbSNP: the NCBI database of genetic variation.* Nucleic Acids Res, 2001. **29**(1): p. 308-11.
320. Corces, M.R., et al., *An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues.* Nature Methods, 2017. **14**: p. 959.
321. Genomics, X., *Chromium Single Cell ATAC Reagent Kits User Guide (v1.1 Chemistry).* 2021.
322. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
323. Neph, S., et al., *BEDOPS: high-performance genomic feature operations.* Bioinformatics, 2012. **28**(14): p. 1919-1920.
324. Danese, A., et al., *EpiScanpy: integrated single-cell epigenomic analysis.* Nature Communications, 2021. **12**(1): p. 5228.
325. Ziegenhain, C., et al., *Comparative Analysis of Single-Cell RNA Sequencing Methods.* Mol Cell, 2017. **65**(4): p. 631-643.e4.
326. Cullen, J.M. and M.J. Stalker, *Chapter 2 - Liver and Biliary System*, in *Jubb, Kennedy & Palmer's Pathology of Domestic Animals: Volume 2 (Sixth Edition)*, M.G. Maxie, Editor. 2016, W.B. Saunders. p. 258-352.e1.
327. Hall, A., et al., *Transaminase abnormalities and adaptations of the liver lobule manifest at specific cut-offs of steatosis.* Scientific Reports, 2017. **7**(1): p. 40977.
328. Deschênes, J., J.-P. Valet, and N. Marceau, *The relationship between cell volume, ploidy, and functional activity in differentiating hepatocytes.* Cell Biophysics, 1981. **3**: p. 321-334.
329. Martin, N.C., et al., *Functional analysis of mouse hepatocytes differing in DNA content: Volume, receptor expression, and effect of IFNγ.* Journal of Cellular Physiology, 2002. **191**(2): p. 138-144.
330. Padovan-Merhar, O., et al., *Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms.* Mol Cell, 2015. **58**(2): p. 339-52.
331. Gentric, G. and C. Desdouets, *Polyploidization in Liver Tissue.* The American Journal of Pathology, 2014. **184**(2): p. 322-331.
332. Batiuk, M.Y., et al., *Identification of region-specific astrocyte subtypes at single cell resolution.* Nature Communications, 2020. **11**(1): p. 1220.
333. Thrupp, N., et al., *Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans.* Cell Rep, 2020. **32**(13): p. 108189.
334. L. Lun, A.T., K. Bach, and J.C. Marioni, *Pooling across cells to normalize single-cell RNA sequencing data with many zero counts.* Genome Biology, 2016. **17**(1): p. 75.
335. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-127.
336. Zhang, M.J., V. Ntranos, and D. Tse, *Determining sequencing depth in a single-cell RNA-seq experiment.* Nature Communications, 2020. **11**(1): p. 774.

337. Söllner, T., et al., *A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion.* Cell, 1993. **75**(3): p. 409-18.
338. de Vries, G.J., A.H. Veenema, and C.H. Brown, *Vasopressin and oxytocin: keys to understanding the neural control of physiology and behaviour.* J Neuroendocrinol, 2012. **24**(4): p. 527.
339. Appleyard, S.M., et al., *Proopiomelanocortin neurons in nucleus tractus solitarius are activated by visceral afferents: regulation by cholecystokinin and opioids.* Journal of Neuroscience, 2005. **25**(14): p. 3578-3585.
340. Souza, L., A.J. Gayban, and Y. Earley, *Tyrosine Hydroxylase-Positive Neurons in the Hypothalamic Paraventricular Nucleus Modulate Glucose Homeostasis in Mice.* The FASEB Journal, 2021. **35**(S1).
341. Kandasamy, P., et al., *Amino acid transporters revisited: New views in health and disease.* Trends in Biochemical Sciences, 2018. **43**(10): p. 752-789.
342. Torgner, I. and E. Kvamme, *Synthesis of transmitter glutamate and the glial-neuron interrelationship.* Molecular and chemical neuropathology, 1990. **12**(1): p. 11-17.
343. Taylor, P. and J.H. Brown, *Synthesis, storage and release of acetylcholine*, in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects. 6th edition*. 1999, Lippincott-Raven.
344. Dicken, M.S., A.R. Hughes, and S.T. Hentges, *Gad1 mRNA as a reliable indicator of altered GABA release from orexigenic neurons in the hypothalamus.* Eur J Neurosci, 2015. **42**(9): p. 2644-53.
345. Donato, R., *S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles.* Int J Biochem Cell Biol, 2001. **33**(7): p. 637-68.
346. Savchenko, V.L., et al., *Microglia and astrocytes in the adult rat brain: comparative immunocytochemical analysis demonstrates the efficacy of lipocortin 1 immunoreactivity.* Neuroscience, 2000. **96**(1): p. 195-203.
347. De Rop, F.V., et al., *Hydrop enables droplet-based single-cell ATAC-seq and single-cell RNA-seq using dissolvable hydrogel beads.* eLife, 2022. **11**: p. e73971.
348. Buenrostro, J.D., et al., *Single-cell chromatin accessibility reveals principles of regulatory variation.* Nature, 2015. **523**: p. 486.
349. Richmond, T.J. and C.A. Davey, *The structure of DNA in the nucleosome core.* Nature, 2003. **423**(6936): p. 145-50.
350. Yoshinari, K., et al., *Omeprazole transactivates human CYP1A1 and CYP1A2 expression through the common regulatory region containing multiple xenobiotic-responsive elements.* Biochemical Pharmacology, 2008. **76**(1): p. 139-145.
351. Wahlländer, A., S. Mohr, and G. Paumgartner, *Assessment of hepatic function. Comparison of caffeine clearance in serum and saliva during the day and at night.* J Hepatol, 1990. **10**(2): p. 129-37.
352. Rulcova, A., et al., *Stereoselective interactions of warfarin enantiomers with the pregnane X nuclear receptor in gene regulation of major drug-metabolizing cytochrome P450 enzymes.* Journal of Thrombosis and Haemostasis, 2010. **8**(12): p. 2708-2717.
353. Hoen, P.A., et al., *Midazolam is a phenobarbital-like cytochrome p450 inducer in rats.* J Pharmacol Exp Ther, 2001. **299**(3): p. 921-7.
354. Greco, D., et al., *Gene expression in human NAFLD.* Am J Physiol Gastrointest Liver Physiol, 2008. **294**(5): p. G1281-7.
355. Ægidius, H.M., et al., *Multi-omics characterization of a diet-induced obese model of non-alcoholic steatohepatitis.* Scientific Reports, 2020. **10**(1): p. 1148.

356. Cohen, J.C., J.D. Horton, and H.H. Hobbs, *Human fatty liver disease: old questions and new insights.* Science, 2011. **332**(6037): p. 1519-23.
357. Musso, G., R. Gambino, and M. Cassader, *Recent insights into hepatic lipid metabolism in non-alcoholic fatty liver disease (NAFLD).* Progress in Lipid Research, 2009. **48**(1): p. 1-26.
358. Pan, X., et al., *Chemokines in Non-alcoholic Fatty Liver Disease: A Systematic Review and Network Meta-Analysis.* Frontiers in immunology, 2020. **11**: p. 1802-1802.
359. Chiba, T., et al., *Diet-induced non-alcoholic fatty liver disease affects expression of major cytochrome P450 genes in a mouse model.* J Pharm Pharmacol, 2016. **68**(12): p. 1567-1576.
360. Donato, M.T., et al., *High-Content Imaging Technology for the Evaluation of Drug-Induced Steatosis Using a Multiparametric Cell-Based Assay.* Journal of Biomolecular Screening, 2011. **17**(3): p. 394-400.
361. Yao, H.R., et al., *Lipotoxicity in HepG2 cells triggered by free fatty acids.* Am J Transl Res, 2011. **3**(3): p. 284-91.
362. Moravcová, A., et al., *The effect of oleic and palmitic acid on induction of steatosis and cytotoxicity on rat hepatocytes in primary culture.* Physiol Res, 2015. **64**(Suppl 5): p. S627-36.
363. Puri, P., et al., *A lipidomic analysis of nonalcoholic fatty liver disease.* Hepatology, 2007. **46**(4): p. 1081-1090.
364. Grishko, V., et al., *Involvement of mtDNA damage in free fatty acid-induced apoptosis.* Free Radic Biol Med, 2005. **38**(6): p. 755-62.
365. Chen, X., et al., *Oleic acid protects saturated fatty acid mediated lipotoxicity in hepatocytes and rat of non-alcoholic steatohepatitis.* Life Sciences, 2018. **203**: p. 291-304.
366. Tolosa, L., et al., *HepG2 cells simultaneously expressing five P450 enzymes for the screening of hepatotoxicity: identification of bioactivable drugs and the potential mechanism of toxicity involved.* Arch Toxicol, 2013. **87**(6): p. 1115-27.
367. Anthérieu, S., et al., *Optimization of the HepaRG cell model for drug metabolism and toxicity studies.* Toxicol In Vitro, 2012. **26**(8): p. 1278-85.
368. Wu, Y., et al., *The HepaRG cell line, a superior in vitro model to L-02, HepG2 and hiHeps cell lines for assessing drug-induced liver injury.* Cell Biol Toxicol, 2016. **32**(1): p. 37-59.
369. Najafi, M., *How to Analyze Real Time qPCR Data?* Biochemistry and Physiology:open access, 2013.
370. Ruijter, J.M., et al., *Efficiency Correction Is Required for Accurate Quantitative PCR Analysis and Reporting.* Clinical Chemistry, 2021. **67**(6): p. 829-842.
371. Zhang, K., et al., *A single-cell atlas of chromatin accessibility in the human genome.* Cell, 2021. **184**(24): p. 5985-6001.e19.
372. Deng, Y., et al., *Spatial profiling of chromatin accessibility in mouse and human tissues.* Nature, 2022. **609**(7926): p. 375-383.
373. Gómez-Lechón, M.J., et al., *Human hepatocytes as a tool for studying toxicity and drug metabolism.* Curr Drug Metab, 2003. **4**(4): p. 292-312.
374. Gómez-Lechón, M.J., et al., *Human hepatocytes in primary culture: the choice to investigate drug metabolism in man.* Curr Drug Metab, 2004. **5**(5): p. 443-62.
375. Grandi, F.C., et al., *Chromatin accessibility profiling by ATAC-seq.* Nat Protoc, 2022. **17**(6): p. 1518-1552.
376. Stuart, T., et al., *Single-cell chromatin state analysis with Signac.* Nature Methods, 2021. **18**(11): p. 1333-1341.
377. Teif, V.B., et al., *Genome-wide nucleosome positioning during embryonic stem cell development.* Nature Structural & Molecular Biology, 2012. **19**(11): p. 1185-1192.
378. Klemm, S.L., Z. Shipony, and W.J. Greenleaf, *Chromatin accessibility and the regulatory epigenome.* Nature Reviews Genetics, 2019. **20**(4): p. 207-220.

379. Yadav, M.L. and B. Mohapatra, *Intergenic*, in *Encyclopedia of Animal Cognition and Behavior*, J. Vonk and T. Shackelford, Editors. 2017, Springer International Publishing: Cham. p. 1-5.
380. Institute, N.H.G.R., *Intergenic Regions.* 2022.
381. Horst, A.K., et al., *CEACAM1 in Liver Injury, Metabolic and Immune Regulation.* Int J Mol Sci, 2018. **19**(10).
382. Ma, L., et al., *α-SNAP Enhances SNARE Zippering by Stabilizing the SNARE Four-Helix Bundle.* Cell Rep, 2016. **15**(3): p. 531-539.
383. Zmuda, E.J., et al., *The Roles of ATF3, an Adaptive-Response Gene, in High-Fat-Diet-Induced Diabetes and Pancreatic β-Cell Dysfunction.* Molecular Endocrinology, 2010. **24**(7): p. 1423-1433.
384. Suganami, T., et al., *Activating transcription factor 3 constitutes a negative feedback mechanism that attenuates saturated Fatty acid/toll-like receptor 4 signaling and macrophage activation in obese adipose tissue.* Circ Res, 2009. **105**(1): p. 25-32.
385. Zelcer, N., et al., *Characterization of drug transport by the human multidrug resistance protein 3 (ABCC3).* J Biol Chem, 2001. **276**(49): p. 46400-7.
386. Castro, A., et al., *APOH is increased in the plasma and liver of type 2 diabetic patients with metabolic syndrome.* Atherosclerosis, 2010. **209**(1): p. 201-5.
387. Dang, T.T.H. and J.W. Yun, *Cytochrome P450 2E1 (CYP2E1) positively regulates lipid catabolism and induces browning in 3T3-L1 white adipocytes.* Life Sci, 2021. **278**: p. 119648.
388. Ni, Q., et al., *Deletion of HNF1α in hepatocytes results in fatty liver-related hepatocellular carcinoma in mice.* FEBS Lett, 2017. **591**(13): p. 1947-1957.
389. He, J., et al., *Hepatocyte nuclear factor 1A suppresses innate immune response by inducing degradation of TBK1 to inhibit steatohepatitis.* Genes & Diseases, 2023. **10**(4): p. 1596-1612.
390. Zhang, T., et al., *Regulation of cytochrome P450 4F11 expression by liver X receptor alpha.* Int Immunopharmacol, 2021. **90**: p. 107240.
391. Hou, Z. and L.H. Matherly, *Biology of the major facilitative folate transporters SLC19A1 and SLC46A1.* Curr Top Membr, 2014. **73**: p. 175-204.
392. Luteijn, R.D., et al., *SLC19A1 transports immunoreactive cyclic dinucleotides.* Nature, 2019. **573**(7774): p. 434-438.
393. Cano, A., et al., *Impaired Function of Solute Carrier Family 19 Leads to Low Folate Levels and Lipid Droplet Accumulation in Hepatocytes.* Biomedicines, 2023. **11**(2).
394. Bröer, S., *Amino Acid Transporters as Disease Modifiers and Drug Targets.* SLAS Discovery, 2018. **23**(4): p. 303-320.
395. Xia, S., et al., *SLC7A2 deficiency promotes hepatocellular carcinoma progression by enhancing recruitment of myeloid-derived suppressors cells.* Cell Death Dis, 2021. **12**(6): p. 570.
396. Loft, A., et al., *Liver-fibrosis-activated transcriptional networks govern hepatocyte reprogramming and intra-hepatic communication.* Cell Metabolism, 2021. **33**(8): p. 1685-1700.e9.
397. Payen, V.L., et al., *Single-cell RNA sequencing of human liver reveals hepatic stellate cell heterogeneity.* JHEP Reports, 2021. **3**(3): p. 100278.
398. Rocque, B., et al., *Creation of a Single Cell RNASeq Meta-Atlas to Define Human Liver Immune Homeostasis.* Frontiers in Immunology, 2021. **12**.
399. Massalha, H., et al., *A single cell atlas of the human liver tumor microenvironment.* Molecular Systems Biology, 2020. **16**(12): p. e9682.
400. Achour, B., J. Barber, and A. Rostami-Hodjegan, *Expression of hepatic drug-metabolizing cytochrome p450 enzymes and their intercorrelations: a meta-analysis.* Drug Metab Dispos, 2014. **42**(8): p. 1349-56.

401. Andrews, T.S., et al., *Single-Cell, Single-Nucleus, and Spatial RNA Sequencing of the Human Liver Identifies Cholangiocyte and Mesenchymal Heterogeneity.* Hepatology Communications, 2021. **n/a**(n/a).
402. Ekstedt, M., et al., *Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up.* Hepatology, 2015. **61**(5): p. 1547-54.
403. Dulai, P.S., et al., *Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis.* Hepatology, 2017. **65**(5): p. 1557-1565.
404. Younossi, Z., et al., *Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention.* Nat Rev Gastroenterol Hepatol, 2018. **15**(1): p. 11-20.
405. Anstee, Q.M., et al., *From NASH to HCC: current concepts and future challenges.* Nat Rev Gastroenterol Hepatol, 2019. **16**(7): p. 411-428.
406. Hernando-Herraez, I., et al., *Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells.* Nature Communications, 2019. **10**(1): p. 4361.
407. Thakur, A., et al., *Ontogeny of drug-metabolizing enzymes.* Enzyme Kinetics in Drug Metabolism: Fundamentals and Applications, 2021: p. 551-593.
408. Wauthier, V., R.K. Verbeeck, and P.B. Calderon, *The effect of ageing on cytochrome p450 enzymes: consequences for drug biotransformation in the elderly.* Curr Med Chem, 2007. **14**(7): p. 745-57.
409. Nicholson, P., et al., *Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors.* Cell Mol Life Sci, 2010. **67**(5): p. 677-700.
410. Mendell, J.T. and H.C. Dietz, *When the message goes awry: disease-producing mutations that influence mRNA content and performance.* Cell, 2001. **107**(4): p. 411-4.
411. Cao, X., et al., *A comprehensive overview of common polymorphic variants that cause missense mutations in human CYPs and UGTs.* Biomed Pharmacother, 2019. **111**: p. 983-992.
412. Genomics, X., *Long Ranger.*
413. Whirl-Carrillo, M., et al., *An Evidence-Based Framework for Evaluating Pharmacogenomics Knowledge for Personalized Medicine.* Clin Pharmacol Ther, 2021. **110**(3): p. 563-572.
414. Whirl-Carrillo, M., et al., *Pharmacogenomics knowledge for personalized medicine.* Clin Pharmacol Ther, 2012. **92**(4): p. 414-7.
415. Rettie, A.E., et al., *Impaired (S)-warfarin metabolism catalysed by the R144C allelic variant of CYP2C9.* Pharmacogenetics and Genomics, 1994. **4**(1): p. 39-42.
416. Scott, S.A., et al., *Combined CYP2C9, VKORC1 and CYP4F2 frequencies among racial and ethnic groups.* Pharmacogenomics, 2010. **11**(6): p. 781-91.
417. Scott, S.A., et al., *Identification of CYP2C19*4B: pharmacogenetic implications for drug metabolism including clopidogrel responsiveness.* Pharmacogenomics J, 2012. **12**(4): p. 297-305.
418. Martis, S., et al., *Multi-ethnic distribution of clinically relevant CYP2C genotypes and haplotypes.* The pharmacogenomics journal, 2013. **13**(4): p. 369-377.
419. Ensembl. *SNP rs3758581*. 23.08.2022; Available from: http://www.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=10:94842366-94843366;v=rs3758581;vdb=variation;vf=167353274.
420. Dalton, R., et al., *Interrogation of CYP2D6 Structural Variant Alleles Improves the Correlation Between CYP2D6 Genotype and CYP2D6-Mediated Metabolic Activity.* Clin Transl Sci, 2020. **13**(1): p. 147-156.
421. Fonseca, D.J., et al., *Whole-Exome Sequencing in Patients Affected by Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis Reveals New Variants Potentially Contributing to the Phenotype.* Pharmgenomics Pers Med, 2021. **14**: p. 287-299.
422. Wang, X., et al., *Direct Comparative Analyses of 10X Genomics Chromium and Smart-seq2.* Genomics, Proteomics & Bioinformatics, 2021. **19**(2): p. 253-266.

423. Genomics, X., *What is the range of compatible cell sizes?*
424. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nature Methods, 2019. **16**(12): p. 1289-1296.
425. Hewitt, N.J., et al., *Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies.* Drug Metab Rev, 2007. **39**(1): p. 159-234.
426. Gómez-Lechón, M.J., J.V. Castell, and M.T. Donato, *The use of hepatocytes to investigate drug toxicity.* Methods Mol Biol, 2010. **640**: p. 389-415.
427. Soldatow, V.Y., et al., *In vitro models for liver toxicity testing.* Toxicol Res (Camb), 2013. **2**(1): p. 23-39.
428. Vinci, B., et al., *Modular bioreactor for primary human hepatocyte culture: medium flow stimulates expression and activity of detoxification genes.* Biotechnol J, 2011. **6**(5): p. 554-64.
429. Aitken, A.E., T.A. Richardson, and E.T. Morgan, *Regulation of drug-metabolizing enzymes and transporters in inflammation.* Annu Rev Pharmacol Toxicol, 2006. **46**: p. 123-49.
430. Renton, K.W., *Regulation of drug metabolism and disposition during inflammation and infection.* Expert Opinion on Drug Metabolism & Toxicology, 2005. **1**(4): p. 629-640.
431. Morgan, E.T., et al., *Regulation of Drug-Metabolizing Enzymes and Transporters in Infection, Inflammation, and Cancer.* Drug Metabolism and Disposition, 2008. **36**(2): p. 205.
432. Davis, A.P., et al., *Comparative Toxicogenomics Database (CTD): update 2021.* Nucleic Acids Res, 2021. **49**(D1): p. D1138-d1143.
433. Davis, A.P., et al., *A CTD-Pfizer collaboration: manual curation of 88,000 scientific articles text mined for drug-disease and drug-phenotype interactions.* Database (Oxford), 2013. **2013**: p. bat080.
434. Mendel, D.B., et al., *In vivo antitumor activity of SU11248, a novel tyrosine kinase inhibitor targeting vascular endothelial growth factor and platelet-derived growth factor receptors: determination of a pharmacokinetic/pharmacodynamic relationship.* Clin Cancer Res, 2003. **9**(1): p. 327-37.
435. Abrams, T.J., et al., *SU11248 inhibits KIT and platelet-derived growth factor receptor beta in preclinical models of human small cell lung cancer.* Mol Cancer Ther, 2003. **2**(5): p. 471-8.
436. Aoudjehane, L., et al., *Novel defatting strategies reduce lipid accumulation in primary human culture models of liver steatosis.* Disease Models & Mechanisms, 2020. **13**(4): p. dmm042663.
437. Imai, Y., et al., *Reduction of hepatosteatosis and lipid levels by an adipose differentiation-related protein antisense oligonucleotide.* Gastroenterology, 2007. **132**(5): p. 1947-54.
438. Imai, Y., et al., *Effects of perilipin 2 antisense oligonucleotide treatment on hepatic lipid metabolism and gene expression.* Physiol Genomics, 2012. **44**(22): p. 1125-31.
439. Libby, A.E., et al., *Perilipin-2 Deletion Impairs Hepatic Lipid Accumulation by Interfering with Sterol Regulatory Element-binding Protein (SREBP) Activation and Altering the Hepatic Lipidome.* J Biol Chem, 2016. **291**(46): p. 24231-24246.
440. Liu, D., et al., *TNFAIP3 Interacting Protein 3 Overexpression Suppresses Nonalcoholic Steatohepatitis by Blocking TAK1 Activation.* Cell Metabolism, 2020. **31**(4): p. 726-740.e8.
441. Zhang, P., et al., *The deubiquitinating enzyme TNFAIP3 mediates inactivation of hepatic ASK1 and ameliorates nonalcoholic steatohepatitis.* Nature Medicine, 2018. **24**(1): p. 84-94.
442. Ge, S.X., D. Jung, and R. Yao, *ShinyGO: a graphical gene-set enrichment tool for animals and plants.* Bioinformatics, 2020. **36**(8): p. 2628-2629.
443. Bode, J.G., et al., *Hepatic acute phase proteins – Regulation by IL-6- and IL-1-type cytokines involving STAT3 and its crosstalk with NF-κB-dependent signaling.* European Journal of Cell Biology, 2012. **91**(6): p. 496-505.
444. Schulze, R.J., et al., *Hepatic Lipophagy: New Insights into Autophagic Catabolism of Lipid Droplets in the Liver.* Hepatology communications, 2017. **1**(5): p. 359-369.

445. Barbosa, A.D. and S. Siniossoglou, *Function of lipid droplet-organelle interactions in lipid homeostasis.* Biochim Biophys Acta Mol Cell Res, 2017. **1864**(9): p. 1459-1468.
446. Lake, A.D., et al., *The adaptive endoplasmic reticulum stress response to lipotoxicity in progressive human nonalcoholic fatty liver disease.* Toxicol Sci, 2014. **137**(1): p. 26-35.
447. Koo, J.H. and C.Y. Han, *Signaling Nodes Associated with Endoplasmic Reticulum Stress during NAFLD Progression.* Biomolecules, 2021. **11**(2).
448. Pan, X., et al., *Chemokines in Non-alcoholic Fatty Liver Disease: A Systematic Review and Network Meta-Analysis.* Frontiers in Immunology, 2020. **11**(1802).
449. Liu, C., et al., *FGF21 protects against hepatic lipotoxicity and macrophage activation to attenuate fibrogenesis in nonalcoholic steatohepatitis.* eLife, 2023. **12**: p. e83075.
450. Hernando-Herraez, I., et al., *Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells.* Nat Commun, 2019. **10**(1): p. 4361.
451. Harries, L.W., et al., *Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing.* Aging Cell, 2011. **10**(5): p. 868-78.
452. Vermulst, M., et al., *Transcription errors induce proteotoxic stress and shorten cellular lifespan.* Nat Commun, 2015. **6**: p. 8065.
453. Işıldak, U., et al., *Temporal changes in the gene expression heterogeneity during brain development and aging.* Scientific reports, 2020. **10**(1): p. 4080.
454. Acun, A., et al., *Liver donor age affects hepatocyte function through age-dependent changes in decellularized liver matrix.* Biomaterials, 2021. **270**: p. 120689.
455. Li, C.Y., et al., *Age-specific regulation of drug-processing genes in mouse liver by ligands of xenobiotic-sensing transcription factors.* Drug Metabolism and Disposition, 2016. **44**(7): p. 1038-1049.
456. Son, J., et al., *Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway.* Nature, 2013. **496**(7443): p. 101-105.
457. Davies, E.A. and M.S. O'Mahony, *Adverse drug reactions in special populations – the elderly.* British Journal of Clinical Pharmacology, 2015. **80**(4): p. 796-807.
458. Cao, S.X., et al., *Genomic profiling of short- and long-term caloric restriction effects in the liver of aging mice.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10630-5.
459. de Magalhães, J.P., J. Curado, and G.M. Church, *Meta-analysis of age-related gene expression profiles identifies common signatures of aging.* Bioinformatics, 2009. **25**(7): p. 875-81.
460. Lee, J.S., et al., *Meta-analysis of gene expression in the mouse liver reveals biomarkers associated with inflammation increased early during aging.* Mech Ageing Dev, 2012. **133**(7): p. 467-78.
461. Bertolotti, M., et al., *Nonalcoholic fatty liver disease and aging: epidemiology to management.* World J Gastroenterol, 2014. **20**(39): p. 14185-204.
462. Ramani, K., et al., *Mechanism and significance of changes in glutamate-cysteine ligase expression during hepatic fibrogenesis.* J Biol Chem, 2012. **287**(43): p. 36341-55.
463. Liu, D. and Y. Xu, *p53, oxidative stress, and aging.* Antioxid Redox Signal, 2011. **15**(6): p. 1669-78.
464. Wang, M.J., et al., *Reversal of hepatocyte senescence after continuous in vivo cell proliferation.* Hepatology, 2014. **60**(1): p. 349-61.
465. Milbank, E., et al., *Liver lipopolysaccharide binding protein prevents hepatic inflammation in physiological and pathological non-obesogenic conditions.* Pharmacological Research, 2023. **187**: p. 106562.
466. Masnoon, N., et al., *What is polypharmacy? A systematic review of definitions.* BMC Geriatrics, 2017. **17**(1): p. 230.
467. Benesic, A., K. Jalal, and A.L. Gerbes, *Drug-Drug Combinations can Enhance Toxicity as shown by Monocyte-Derived Hepatocyte-like Cells from Patients with Idiosyncratic Drug-Induced Liver Injury.* Toxicol Sci, 2019.

468. Douros, A., et al., *Drug-induced liver injury: results from the hospital-based Berlin Case-Control Surveillance Study.* British journal of clinical pharmacology, 2015. **79**(6): p. 988-999.
469. Juurlink, D.N., et al., *Drug-drug interactions among elderly patients hospitalized for drug toxicity.* Jama, 2003. **289**(13): p. 1652-8.
470. Chen, L., et al., *Transcriptomic profiling of hepatic tissues for drug metabolism genes in nonalcoholic fatty liver disease: A study of human and animals.* Front Endocrinol (Lausanne), 2022. **13**: p. 1034494.
471. Odom, D.T., et al., *Control of pancreas and liver gene expression by HNF transcription factors.* Science, 2004. **303**(5662): p. 1378-81.
472. Xanthopoulos, K.G., et al., *The different tissue transcription patterns of genes for HNF-1, C/EBP, HNF-3, and HNF-4, protein factors that govern liver-specific transcription.* Proc Natl Acad Sci U S A, 1991. **88**(9): p. 3807-11.
473. Hu, P., et al., *Dissecting Cell-Type Composition and Activity-Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq.* Mol Cell, 2017. **68**(5): p. 1006-1015.e7.
474. Lake, B.B., et al., *A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA.* Sci Rep, 2017. **7**(1): p. 6031.
475. Lake, B.B., et al., *Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain.* Science, 2016. **352**(6293): p. 1586-90.
476. Kalish, B.T., et al., *Single-nucleus RNA sequencing of mouse auditory cortex reveals critical period triggers and brakes.* Proc Natl Acad Sci U S A, 2020. **117**(21): p. 11744-11752.
477. Koenitzer, J.R., et al., *Single-Nucleus RNA-Sequencing Profiling of Mouse Lung. Reduced Dissociation Bias and Improved Rare Cell-Type Detection Compared with Single-Cell RNA Sequencing.* Am J Respir Cell Mol Biol, 2020. **63**(6): p. 739-747.
478. Wu, H., et al., *Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis.* J Am Soc Nephrol, 2019. **30**(1): p. 23-32.
479. Wilson, P.C., et al., *The single-cell transcriptomic landscape of early human diabetic nephropathy.* Proc Natl Acad Sci U S A, 2019. **116**(39): p. 19619-19625.
480. Denisenko, E., et al., *Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows.* Genome Biol, 2020. **21**(1): p. 130.
481. Lake, B.B., et al., *A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys.* Nat Commun, 2019. **10**(1): p. 2832.
482. Selewa, A., et al., *Systematic Comparison of High-throughput Single-Cell and Single-Nucleus Transcriptomes during Cardiomyocyte Differentiation.* Sci Rep, 2020. **10**(1): p. 1535.
483. Wolfien, M., et al., *Single-Nucleus Sequencing of an Entire Mammalian Heart: Cell Type Composition and Velocity.* Cells, 2020. **9**(2).
484. Basile, G., et al., *Using single-nucleus RNA-sequencing to interrogate transcriptomic profiles of archived human pancreatic islets.* Genome Medicine, 2021. **13**(1): p. 128.
485. Nault, R., et al., *Single-Nuclei RNA Sequencing Assessment of the Hepatic Effects of 2,3,7,8-Tetrachlorodibenzo-p-dioxin.* Cell Mol Gastroenterol Hepatol, 2021. **11**(1): p. 147-159.
486. Lacar, B., et al., *Nuclear RNA-seq of single neurons reveals molecular signatures of activation.* Nat Commun, 2016. **7**: p. 11022.
487. van den Brink, S.C., et al., *Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations.* Nat Methods, 2017. **14**(10): p. 935-936.
488. Lee, H., et al., *External RNA Controls Consortium Beta Version Update.* Journal of Genomics, 2016. **4**: p. 19-22.
489. Svensson, V., et al., *Power analysis of single-cell RNA-sequencing experiments.* Nature Methods, 2017. **14**(4): p. 381-387.

490. Baran-Gale, J., T. Chandra, and K. Kirschner, *Experimental design for single-cell RNA sequencing.* Brief Funct Genomics, 2018. **17**(4): p. 233-239.
491. Lafzi, A., et al., *Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies.* Nat Protoc, 2018. **13**(12): p. 2742-2757.
492. Tang, F., et al., *mRNA-Seq whole-transcriptome analysis of a single cell.* Nature Methods, 2009. **6**(5): p. 377-382.
493. Sheng, K., et al., *Effective detection of variation in single-cell transcriptomes using MATQ-seq.* Nature Methods, 2017. **14**(3): p. 267-270.
494. Mereu, E., et al., *Benchmarking single-cell RNA-sequencing protocols for cell atlas projects.* Nature Biotechnology, 2020. **38**(6): p. 747-755.
495. Ding, J., et al., *Systematic comparison of single-cell and single-nucleus RNA-sequencing methods.* Nature Biotechnology, 2020. **38**(6): p. 737-746.
496. Zilionis, R., et al., *Single-cell barcoding and sequencing using droplet microfluidics.* Nat Protoc, 2017. **12**(1): p. 44-73.
497. Gierahn, T.M., et al., *Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput.* Nat Methods, 2017. **14**(4): p. 395-398.
498. Hochgerner, H., et al., *STRT-seq-2i: dual-index 5′ single cell and nucleus RNA-seq on an addressable microwell array.* Sci Rep, 2017. **7**(1): p. 16327.
499. Lafzi, A., et al., *Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies.* Nature Protocols, 2018. **13**(12): p. 2742-2757.
500. Krishnaswami, S.R., et al., *Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons.* Nature Protocols, 2016. **11**(3): p. 499-524.
501. Schmoller, K.M. and J.M. Skotheim, *The Biosynthetic Basis of Cell Size Control.* Trends Cell Biol, 2015. **25**(12): p. 793-802.
502. Vargas-Garcia, C.A., K.R. Ghusinga, and A. Singh, *Cell size control and gene expression homeostasis in single-cells.* Current Opinion in Systems Biology, 2018. **8**: p. 109-116.
503. Froemke, R.C. and I. Carcea, *Oxytocin and brain plasticity*, in *Principles of gender-specific medicine*. 2017, Elsevier. p. 161-182.
504. Dowsett, G.K.C., et al., *A survey of the mouse hindbrain in the fed and fasted states using single-nucleus RNA sequencing.* Mol Metab, 2021. **53**: p. 101240.
505. Keren-Shaul, H., et al., *A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease.* Cell, 2017. **169**(7): p. 1276-1290.e17.
506. Zhang, S., et al., *The origins and functions of hepatic polyploidy.* Cell Cycle, 2019. **18**(12): p. 1302-1315.
507. Park, H.-J., et al., *Cellular heterogeneity and plasticity during NAFLD progression.* Frontiers in Molecular Biosciences, 2023. **10**.
508. Corton, J.C., et al., *Sources of variance in baseline gene expression in the rodent liver.* Mutat Res, 2012. **746**(2): p. 104-12.
509. Parent, R., et al., *Origin and characterization of a human bipotent liver progenitor cell line.* Gastroenterology, 2004. **126**(4): p. 1147-56.
510. Kammerer, S. *Three-Dimensional Liver Culture Systems to Maintain Primary Hepatic Properties for Toxicological Analysis In Vitro*. International Journal of Molecular Sciences, 2021. **22**, DOI: 10.3390/ijms221910214.
511. Donato, M.T., G. Gallego-Ferrer, and L. Tolosa, *In Vitro Models for Studying Chronic Drug-Induced Liver Injury.* Int J Mol Sci, 2022. **23**(19).
512. Yokoyama, Y., et al., *Comparison of Drug Metabolism and Its Related Hepatotoxic Effects in HepaRG, Cryopreserved Human Hepatocytes, and HepG2 Cell Cultures.* Biol Pharm Bull, 2018. **41**(5): p. 722-732.

513. Huggett, Z.J., et al., *A Comparison of Primary Human Hepatocytes and Hepatoma Cell Lines to Model the Effects of Fatty Acids, Fructose and Glucose on Liver Cell Lipid Accumulation.* Nutrients, 2022. **15**(1).
514. Tolosa, L., et al., *Advantageous use of HepaRG cells for the screening and mechanistic study of drug-induced steatosis.* Toxicol Appl Pharmacol, 2016. **302**: p. 1-9.
515. Tanner, N., et al., *Regulation of Drug Metabolism by the Interplay of Inflammatory Signaling, Steatosis, and Xeno-Sensing Receptors in HepaRG Cells.* Drug Metab Dispos, 2018. **46**(4): p. 326-335.
516. Pelechá, M., et al. *Cell Models and Omics Techniques for the Study of Nonalcoholic Fatty Liver Disease: Focusing on Stem Cell-Derived Cell Models*. Antioxidants, 2022. **11**, DOI: 10.3390/antiox11010086.
517. Ipsen, D.H., J. Lykkesfeldt, and P. Tveden-Nyborg, *Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease.* Cellular and Molecular Life Sciences, 2018. **75**(18): p. 3313-3327.
518. Bence, K.K. and M.J. Birnbaum, *Metabolic drivers of non-alcoholic fatty liver disease.* Mol Metab, 2021. **50**: p. 101143.
519. Chainuvati, S., et al., *Combined phenotypic assessment of cytochrome p450 1A2, 2C9, 2C19, 2D6, and 3A, N-acetyltransferase-2, and xanthine oxidase activities with the "Cooperstown 5+1 cocktail".* Clin Pharmacol Ther, 2003. **74**(5): p. 437-47.
520. Tanaka, E., N. Kurata, and H. Yasuhara, *How useful is the 'cocktail approach' for evaluating human hepatic drug metabolizing capacity using cytochrome P450 phenotyping probes in vivo?* Journal of Clinical Pharmacy and Therapeutics, 2003. **28**(3): p. 157-165.
521. Spaggiari, D., L. Geiser, and S. Rudaz, *Coupling ultra-high-pressure liquid chromatography with mass spectrometry for in-vitro drug-metabolism studies.* TrAC Trends in Analytical Chemistry, 2014. **63**: p. 129-139.
522. Berger, B., et al., *Comparison Of Liver Cell Models Using The Basel Phenotyping Cocktail.* Frontiers in Pharmacology, 2016. **7**(443).
523. Alshabi, A., et al., *A cocktail probe approach to evaluate the effect of hormones on the expression and activity of CYP enzymes in human hepatocytes with conditions simulating late stage of pregnancy.* Eur J Clin Pharmacol, 2023. **79**(6): p. 815-827.
524. Lammers, L.A., et al., *Effect of Short-Term Fasting on Systemic Cytochrome P450-Mediated Drug Metabolism in Healthy Subjects: A Randomized, Controlled, Crossover Study Using a Cocktail Approach.* Clin Pharmacokinet, 2017. **56**(10): p. 1231-1244.
525. Toda, A., et al., *Effects of aging and rifampicin pretreatment on the pharmacokinetics of human cytochrome P450 probes caffeine, warfarin, omeprazole, metoprolol and midazolam in common marmosets genotyped for cytochrome P450 2C19.* Xenobiotica, 2018. **48**(7): p. 720-726.
526. Mogi, M., et al., *Simultaneous pharmacokinetics assessment of caffeine, warfarin, omeprazole, metoprolol, and midazolam intravenously or orally administered to Microminipigs.* J Toxicol Sci, 2012. **37**(6): p. 1157-64.
527. Ingelman-Sundberg, M., *Pharmacogenetics: an opportunity for a safer and more efficient pharmacotherapy.* J Intern Med, 2001. **250**(3): p. 186-200.
528. Spatzenegger, M. and W. Jaeger, *Clinical importance of hepatic cytochrome P450 in drug metabolism.* Drug metabolism reviews, 1995. **27**(3): p. 397-417.
529. Rao, X., et al., *An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis.* Biostat Bioinforma Biomath, 2013. **3**(3): p. 71-85.
530. Arocho, A., et al., *Validation of the 2-DeltaDeltaCt calculation as an alternate method of data analysis for quantitative PCR of BCR-ABL P210 transcripts.* Diagn Mol Pathol, 2006. **15**(1): p. 56-61.

531. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
532. Svec, D., et al., *How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments.* Biomol Detect Quantif, 2015. **3**: p. 9-16.
533. Anthérieu, S., et al., *Induction of vesicular steatosis by amiodarone and tetracycline is associated with up-regulation of lipogenic genes in heparg cells.* Hepatology, 2011. **53**(6): p. 1895-1905.
534. Pant, A., E.A. Rondini, and T.A. Kocarek, *Farnesol induces fatty acid oxidation and decreases triglyceride accumulation in steatotic HepaRG cells.* Toxicol Appl Pharmacol, 2019. **365**: p. 61-70.
535. Jamwal, R. and B.J. Barlock, *Nonalcoholic Fatty Liver Disease (NAFLD) and Hepatic Cytochrome P450 (CYP) Enzymes.* Pharmaceuticals (Basel), 2020. **13**(9).
536. Jackson, J.P., et al., *Contextualizing Hepatocyte Functionality of Cryopreserved HepaRG Cell Cultures.* Drug Metab Dispos, 2016. **44**(9): p. 1463-79.
537. Martínez-Jiménez, C.P., et al., *Underexpressed Coactivators PGC1α AND SRC1 Impair Hepatocyte Nuclear Factor 4α Function and Promote Dedifferentiation in Human Hepatoma Cells *.* Journal of Biological Chemistry, 2006. **281**(40): p. 29840-29849.
538. Hammour, M.M., et al., *Optimisation of the HepaRG cell line model for drug toxicity studies using two different cultivation conditions: advantages and limitations.* Archives of Toxicology, 2022. **96**(9): p. 2511-2521.
539. Yadav, J., et al., *Recent developments in in vitro and in vivo models for improved translation of preclinical pharmacokinetics and pharmacodynamics data.* Drug Metab Rev, 2021. **53**(2): p. 207-233.
540. Bhogal, R.H., et al., *Isolation of primary human hepatocytes from normal and diseased liver tissue: a one hundred liver experience.* PLoS One, 2011. **6**(3): p. e18222.
541. Wilkening, S., F. Stahl, and A. Bader, *Comparison of primary human hepatocytes and hepatoma cell line Hepg2 with regard to their biotransformation properties.* Drug Metab Dispos, 2003. **31**(8): p. 1035-42.
542. Schwartz, R.E., et al., *Pluripotent stem cell-derived hepatocyte-like cells.* Biotechnology Advances, 2014. **32**(2): p. 504-513.
543. Guo, L., et al., *Similarities and differences in the expression of drug-metabolizing enzymes between human hepatic cell lines and primary human hepatocytes.* Drug Metab Dispos, 2011. **39**(3): p. 528-38.
544. Jetten, M.J., et al., *Baseline and genotoxic compound induced gene expression profiles in HepG2 and HepaRG compared to primary human hepatocytes.* Toxicol In Vitro, 2013. **27**(7): p. 2031-40.
545. Elaut, G., et al., *Molecular Mechanisms Underlying the Dedifferentiation Process of Isolated Hepatocytes and Their Cultures.* Current Drug Metabolism, 2006. **7**(6): p. 629-660.
546. Godoy, P., et al., *Extracellular matrix modulates sensitivity of hepatocytes to fibroblastoid dedifferentiation and transforming growth factor β–induced apoptosis.* Hepatology, 2009. **49**(6): p. 2031-2043.
547. Schyschka, L., et al., *Hepatic 3D cultures but not 2D cultures preserve specific transporter activity for acetaminophen-induced hepatotoxicity.* Archives of Toxicology, 2013. **87**(8): p. 1581-1593.
548. Gómez-Lechón, M.J., et al., *Competency of different cell models to predict human hepatotoxic drugs.* Expert Opin Drug Metab Toxicol, 2014. **10**(11): p. 1553-68.
549. LeCluyse, E.L., et al., *Isolation and culture of primary human hepatocytes.* Methods Mol Biol, 2005. **290**: p. 207-29.

550. Handin, N., et al., *Conditions for maintenance of hepatocyte differentiation and function in 3D cultures.* iScience, 2021. **24**(11): p. 103235.
551. Mansisidor, A.R. and V.I. Risca, *Chromatin accessibility: methods, mechanisms, and biological insights.* Nucleus, 2022. **13**(1): p. 236-276.
552. Klemm, S.L., Z. Shipony, and W.J. Greenleaf, *Chromatin accessibility and the regulatory epigenome.* Nat Rev Genet, 2019. **20**(4): p. 207-220.
553. Granja, J.M., et al., *Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia.* Nature Biotechnology, 2019. **37**(12): p. 1458-1465.
554. Pervolarakis, N., et al., *Integrated Single-Cell Transcriptomics and Chromatin Accessibility Analysis Reveals Regulators of Mammary Epithelial Cell Identity.* Cell Rep, 2020. **33**(3): p. 108273.
555. Denyer, T. and M.C.P. Timmermans, *High-throughput single-cell RNA sequencing.* Trends in Plant Science, 2022. **27**(1): p. 104-105.
556. Yan, F., et al., *From reads to insight: a hitchhiker's guide to ATAC-seq data analysis.* Genome Biology, 2020. **21**(1): p. 22.
557. Luecken, M.D., et al., *Benchmarking atlas-level data integration in single-cell genomics.* Nat Methods, 2022. **19**(1): p. 41-50.
558. Tran, H.T.N., et al., *A benchmark of batch-effect correction methods for single-cell RNA sequencing data.* Genome Biol, 2020. **21**(1): p. 12.
559. Büttner, M., et al., *A test metric for assessing single-cell RNA-seq batch correction.* Nat Methods, 2019. **16**(1): p. 43-49.
560. Chen, W., et al., *A multicenter study benchmarking single-cell RNA sequencing technologies using reference samples.* Nature Biotechnology, 2021. **39**(9): p. 1103-1114.
561. Chen, H., et al., *Assessment of computational methods for the analysis of single-cell ATAC-seq data.* Genome Biology, 2019. **20**(1): p. 241.
562. Luecken, M.D. and F.J. Theis, *Current best practices in single-cell RNA-seq analysis: a tutorial.* Molecular Systems Biology, 2019. **15**(6): p. e8746.
563. Lyu, P., et al., *Gene regulatory networks controlling temporal patterning, neurogenesis, and cell-fate specification in mammalian retina.* Cell Rep, 2021. **37**(7): p. 109994.
564. Miller, A. and W.W. Jedrzejczak, *[Albumin--biological functions and clinical significance].* Postepy Hig Med Dosw, 2001. **55**(1): p. 17-36.
565. Allen-Jennings, A.E., et al., *The roles of ATF3 in liver dysfunction and the regulation of phosphoenolpyruvate carboxykinase gene expression.* J Biol Chem, 2002. **277**(22): p. 20020-5.
566. Huang, K.-W., et al., *Liver Activation of Hepatocellular Nuclear Factor-4α by Small Activating RNA Rescues Dyslipidemia and Improves Metabolic Profile.* Molecular Therapy - Nucleic Acids, 2020. **19**: p. 361-370.
567. Lu, H., *Crosstalk of HNF4α with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids.* Acta Pharm Sin B, 2016. **6**(5): p. 393-408.
568. Yi, M., et al., *Functional characterization of a common CYP4F11 genetic variant and identification of functionally defective CYP4F11 variants in erythromycin metabolism and 20-HETE synthesis.* Archives of Biochemistry and Biophysics, 2017. **620**: p. 43-51.
569. Das, A., et al., *CYP2J2 Molecular Recognition: A New Axis for Therapeutic Design.* Pharmacology & Therapeutics, 2020. **215**: p. 107601.
570. Chen, G., et al., *CYP2J2 overexpression attenuates nonalcoholic fatty liver disease induced by high-fat diet in mice.* Am J Physiol Endocrinol Metab, 2015. **308**(2): p. E97-e110.
571. Hegenbarth, J.-C., et al., *Perspectives on Bulk-Tissue RNA Sequencing and Single-Cell RNA Sequencing for Cardiac Transcriptomics.* Frontiers in Molecular Medicine, 2022. **2**.

572. Yu, X., et al., *Statistical and Bioinformatics Analysis of Data from Bulk and Single-Cell RNA Sequencing Experiments.* Methods Mol Biol, 2021. **2194**: p. 143-175.
573. Li, X. and C.-Y. Wang, *From bulk, single-cell to spatial RNA sequencing.* International Journal of Oral Science, 2021. **13**(1): p. 36.
574. Vinci, B., et al., *Modular bioreactor for primary human hepatocyte culture: medium flow stimulates expression and activity of detoxification genes.* Biotechnology journal, 2011. **6**(5): p. 554-564.
575. Hayhurst, G.P., et al., *Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis.* Mol Cell Biol, 2001. **21**(4): p. 1393-403.
576. Odom, D.T., et al., *Tissue-specific transcriptional regulation has diverged significantly between human and mouse.* Nat Genet, 2007. **39**(6): p. 730-2.
577. Sanchez-Quant E., Richter M.L., Colomé-Tatché M., Martinez-Jimenez C.P., *Single-cell metabolic profiling reveals subgroups of primary human hepatocytes showing heterogeneous responses to a drug challenge.* ArrayExpress, 2022.
578. Sun, D., et al., *Why 90% of clinical drug development fails and how to improve it?* Acta Pharm Sin B, 2022. **12**(7): p. 3049-3062.
579. Peter, J.B., et al., *Barriers to new drug development in respiratory disease.* European Respiratory Journal, 2015. **45**(5): p. 1197.
580. Guengerich, F.P., *Mechanisms of drug toxicity and relevance to pharmaceutical development.* Drug Metab Pharmacokinet, 2011. **26**(1): p. 3-14.
581. Pognan, F., et al., *The evolving role of investigative toxicology in the pharmaceutical industry.* Nature Reviews Drug Discovery, 2023. **22**(4): p. 317-335.
582. Dowden, H. and J. Munro, *Trends in clinical success rates and therapeutic focus.* Nat Rev Drug Discov, 2019. **18**(7): p. 495-496.
583. Harrison, R.K., *Phase II and phase III failures: 2013-2015.* Nat Rev Drug Discov, 2016. **15**(12): p. 817-818.
584. Okada, D., C. Zheng, and J.H. Cheng, *Mathematical model for the relationship between single-cell and bulk gene expression to clarify the interpretation of bulk gene expression data.* Computational and Structural Biotechnology Journal, 2022. **20**: p. 4850-4859.
585. Timm, A. and J.M. Kolesar, *Crizotinib for the treatment of non-small-cell lung cancer.* Am J Health Syst Pharm, 2013. **70**(11): p. 943-7.
586. Katarey, D. and S. Verma, *Drug-induced liver injury.* Clin Med (Lond), 2016. **16**(Suppl 6): p. s104-s109.
587. Kuna, L., et al., *Models of Drug Induced Liver Injury (DILI) - Current Issues and Future Perspectives.* Curr Drug Metab, 2018. **19**(10): p. 830-838.
588. Chen, M., et al., *Drug-induced liver injury: Interactions between drug properties and host factors.* J Hepatol, 2015. **63**(2): p. 503-14.
589. Babai, S., L. Auclert, and H. Le-Louët, *Safety data and withdrawal of hepatotoxic drugs.* Therapie, 2021. **76**(6): p. 715-723.
590. Weaver, R.J., et al., *Managing the challenge of drug-induced liver injury: a roadmap for the development and deployment of preclinical predictive models.* Nat Rev Drug Discov, 2020. **19**(2): p. 131-148.
591. Takebe, T., R. Imai, and S. Ono, *The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development.* Clinical and Translational Science, 2018. **11**(6): p. 597-606.
592. Sun, D., et al., *Why 90% of clinical drug development fails and how to improve it?* Acta Pharmaceutica Sinica B, 2022. **12**(7): p. 3049-3062.

593. Arellano, A.L., et al., *Multiple adverse drug reactions and genetic polymorphism testing: A case report with negative result.* Medicine (Baltimore), 2017. **96**(45): p. e8505.
594. Zhang, L.L., et al., *Genetic polymorphisms affect efficacy and adverse drug reactions of DMARDs in rheumatoid arthritis.* Pharmacogenet Genomics, 2014. **24**(11): p. 531-8.
595. Moore, L.D., T. Le, and G. Fan, *DNA Methylation and Its Basic Function.* Neuropsychopharmacology, 2013. **38**(1): p. 23-38.
596. Tsompana, M. and M.J. Buck, *Chromatin accessibility: a window into the genome.* Epigenetics & Chromatin, 2014. **7**(1): p. 33.
597. Luquette, L.J., et al., *Identification of somatic mutations in single cell DNA-seq using a spatial model of allelic imbalance.* Nature Communications, 2019. **10**(1): p. 3908.
598. Fan, X., et al., *SMOOTH-seq: single-cell genome sequencing of human cells on a third-generation sequencing platform.* Genome Biology, 2021. **22**(1): p. 195.
599. Probst, V., et al., *Benchmarking full-length transcript single cell mRNA sequencing protocols.* BMC Genomics, 2022. **23**(1): p. 860.
600. Sim, S.C. and M. Ingelman-Sundberg, *The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects.* Hum Genomics, 2010. **4**(4): p. 278-81.
601. Goldstein, J.A., *Clinical relevance of genetic polymorphisms in the human CYP2C subfamily.* Br J Clin Pharmacol, 2001. **52**(4): p. 349-55.
602. Ferguson, R.J., et al., *A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin.* J Pharmacol Exp Ther, 1998. **284**(1): p. 356-61.
603. Clermont, V., et al., *Activity and mRNA expression levels of selected cytochromes P450 in various sections of the human small intestine.* British Journal of Clinical Pharmacology, 2019. **85**(6): p. 1367-1377.
604. Sanderson, S., J. Emery, and J. Higgins, *CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: a HuGEnet systematic review and meta-analysis.* Genet Med, 2005. **7**(2): p. 97-104.
605. Schalekamp, T., et al., *Effects of cytochrome P450 2C9 polymorphisms on phenprocoumon anticoagulation status.* Clin Pharmacol Ther, 2004. **76**(5): p. 409-17.
606. Gonzalez-Covarrubias, V., et al., *Pharmacogenetic Variation in Over 100 Genes in Patients Receiving Acenocumarol.* Front Pharmacol, 2017. **8**: p. 863.
607. Holstein, A., et al., *Association between CYP2C9 slow metabolizer genotypes and severe hypoglycaemia on medication with sulphonylurea hypoglycaemic agents.* Br J Clin Pharmacol, 2005. **60**(1): p. 103-6.
608. He, S.-M., et al., *Clinical drugs undergoing polymorphic metabolism by human cytochrome P450 2C9 and the implication in drug development.* Current medicinal chemistry, 2011. **18**(5): p. 667-713.
609. Konstandi, M. and E.O. Johnson, *Age-related modifications in CYP-dependent drug metabolism: role of stress.* Frontiers in Endocrinology, 2023. **14**.
610. Parkinson, A., et al., *The effects of gender, age, ethnicity, and liver cirrhosis on cytochrome P450 enzyme activity in human liver microsomes and inducibility in cultured human hepatocytes.* Toxicol Appl Pharmacol, 2004. **199**(3): p. 193-209.
611. Schwartz, J.B., *The Current State of Knowledge on Age, Sex, and Their Interactions on Clinical Pharmacology.* Clinical Pharmacology & Therapeutics, 2007. **82**(1): p. 87-96.
612. Osanlou, R., et al., *Adverse drug reactions, multimorbidity and polypharmacy: a prospective analysis of 1 month of medical admissions.* BMJ Open, 2022. **12**(7): p. e055551.
613. Dagli, R.J. and A. Sharma, *Polypharmacy: a global risk factor for elderly people.* J Int Oral Health, 2014. **6**(6): p. i-ii.

614. Lucena, M.I., et al., *Drug-induced liver injury in older people.* The Lancet Gastroenterology & Hepatology, 2020. **5**(9): p. 862-874.
615. Lavan, A.H. and P. Gallagher, *Predicting risk of adverse drug reactions in older adults.* Therapeutic advances in drug safety, 2016. **7**(1): p. 11-22.
616. Doan, J., et al., *Prevalence and risk of potential cytochrome P450-mediated drug-drug interactions in older hospitalized patients with polypharmacy.* Ann Pharmacother, 2013. **47**(3): p. 324-32.
617. Dubrall, D., et al., *Adverse drug reactions in older adults: a retrospective comparative analysis of spontaneous reports to the German Federal Institute for Drugs and Medical Devices.* BMC Pharmacology and Toxicology, 2020. **21**(1): p. 25.
618. Cherubini, A., et al., *Fighting against age discrimination in clinical trials.* J Am Geriatr Soc, 2010. **58**(9): p. 1791-6.
619. van Marum, R.J., *Underrepresentation of the elderly in clinical trials, time for action.* Br J Clin Pharmacol, 2020. **86**(10): p. 2014-2016.
620. van Deursen, J.M., *The role of senescent cells in ageing.* Nature, 2014. **509**(7501): p. 439-46.
621. Kuilman, T., et al., *The essence of senescence.* Genes Dev, 2010. **24**(22): p. 2463-79.
622. Rodier, F. and J. Campisi, *Four faces of cellular senescence.* J Cell Biol, 2011. **192**(4): p. 547-56.
623. Kregel, K.C. and H.J. Zhang, *An integrated view of oxidative stress in aging: basic mechanisms, functional effects, and pathological considerations.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2007. **292**(1): p. R18-R36.
624. Bárcena, B., et al., *Aging Induces Hepatic Oxidative Stress and Nuclear Proteomic Remodeling in Liver from Wistar Rats.* Antioxidants (Basel), 2021. **10**(10).
625. Poulose, N. and R. Raju, *Aging and injury: alterations in cellular energetics and organ function.* Aging Dis, 2014. **5**(2): p. 101-8.
626. Radonjić, T., et al., *Aging of Liver in Its Different Diseases.* Int J Mol Sci, 2022. **23**(21).
627. Liu, F.-T. and G.A. Rabinovich, *Galectins as modulators of tumour progression.* Nature Reviews Cancer, 2005. **5**(1): p. 29-41.
628. Baek, J.-H., et al., *Galectin-1 accelerates high-fat diet-induced obesity by activation of peroxisome proliferator-activated receptor gamma (PPARγ) in mice.* Cell Death & Disease, 2021. **12**(1): p. 66.
629. Lu, S.C., *Regulation of glutathione synthesis.* Mol Aspects Med, 2009. **30**(1-2): p. 42-59.
630. Franceschi, C., et al., *Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans.* Mech Ageing Dev, 2007. **128**(1): p. 92-105.
631. Salama, R., et al., *Cellular senescence and its effector programs.* Genes Dev, 2014. **28**(2): p. 99-114.
632. Saavedra, D., et al., *Aging and chronic inflammation: highlights from a multidisciplinary workshop.* Immunity & Ageing, 2023. **20**(1): p. 25.
633. Mendenhall, A.R., et al., *Cell-to-cell variation in gene expression and the aging process.* Geroscience, 2021. **43**(1): p. 181-196.
634. Yamamoto, R., et al., *Tissue-specific impacts of aging and genetics on gene expression patterns in humans.* Nature Communications, 2022. **13**(1): p. 5803.
635. Wang, S., et al., *Single-Cell Transcriptomic Atlas of Primate Ovarian Aging.* Cell, 2020. **180**(3): p. 585-600.e19.
636. Kolodziejczyk, A.A., et al., *Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation.* Cell Stem Cell, 2015. **17**(4): p. 471-85.
637. Bahar Halpern, K., et al., *Bursty gene expression in the intact mammalian liver.* Mol Cell, 2015. **58**(1): p. 147-56.
638. Dodd, K.S., et al., *Exclusion of older adults and women from recent trials of acute coronary syndromes.* J Am Geriatr Soc, 2011. **59**(3): p. 506-11.

639. McMurdo, M.E., M.D. Witham, and N.D. Gillespie, *Including older people in clinical research.* Bmj, 2005. **331**(7524): p. 1036-7.
640. Drenth-van Maanen, A.C., I. Wilting, and P.A.F. Jansen, *Prescribing medicines to older people—How to consider the impact of ageing on human organ and body functions.* British Journal of Clinical Pharmacology, 2020. **86**(10): p. 1921-1930.
641. Tannenbaum, C. and D. Day, *Age and sex in drug development and testing for adults.* Pharmacological Research, 2017. **121**: p. 83-93.
642. Govaere, O., et al., *A proteo-transcriptomic map of non-alcoholic fatty liver disease signatures.* Nature Metabolism, 2023. **5**(4): p. 572-578.
643. Santhekadur, P.K., D.P. Kumar, and A.J. Sanyal, *Preclinical models of non-alcoholic fatty liver disease.* J Hepatol, 2018. **68**(2): p. 230-237.
644. Tarantino, G., et al., *A prospective study of acute drug-induced liver injury in patients suffering from non-alcoholic fatty liver disease.* Hepatol Res, 2007. **37**(6): p. 410-5.
645. Donato, M. and L. Tolosa, *High-Content Screening for the Detection of Drug-Induced Oxidative Stress in Liver Cells.* Antioxidants (Basel), 2021. **10**(1).
646. Gao, H., et al., *CYP4A11 is involved in the development of nonalcoholic fatty liver disease via ROS-induced lipid peroxidation and inflammation.* Int J Mol Med, 2020. **45**(4): p. 1121-1129.
647. Wang, M. and R.J. Kaufman, *Protein misfolding in the endoplasmic reticulum as a conduit to human disease.* Nature, 2016. **529**(7586): p. 326-335.
648. Dubois, V., et al., *Endoplasmic reticulum stress actively suppresses hepatic molecular identity in damaged liver.* Molecular Systems Biology, 2020. **16**(5): p. e9156.
649. Kim, K.H. and M.S. Lee, *GDF15 as a central mediator for integrated stress response and a promising therapeutic molecule for metabolic disorders and NASH.* Biochim Biophys Acta Gen Subj, 2021. **1865**(3): p. 129834.
650. Han, C.Y., et al., *PHLDA3 overexpression in hepatocytes by endoplasmic reticulum stress via IRE1-Xbp1s pathway expedites liver injury.* Gut, 2016. **65**(8): p. 1377-88.
651. Lork, M., K. Verhelst, and R. Beyaert, *CYLD, A20 and OTULIN deubiquitinases in NF-κB signaling and cell death: so similar, yet so different.* Cell Death Differ, 2017. **24**(7): p. 1172-1183.
652. Pakos-Zebrucka, K., et al., *The integrated stress response.* EMBO reports, 2016. **17**(10): p. 1374-1395.
653. Torrence, M.E., et al., *The mTORC1-mediated activation of ATF4 promotes protein and glutathione synthesis downstream of growth signals.* eLife, 2021. **10**: p. e63326.
654. Li, K., et al., *Liver-specific Gene Inactivation of the Transcription Factor ATF4 Alleviates Alcoholic Liver Steatosis in Mice.* J Biol Chem, 2016. **291**(35): p. 18536-46.
655. Hong, L., et al., *New role and molecular mechanism of Gadd45a in hepatic fibrosis.* World J Gastroenterol, 2016. **22**(9): p. 2779-88.
656. Ahmad, T., M.A. Valentovic, and G.O. Rankin, *Effects of cytochrome P450 single nucleotide polymorphisms on methadone metabolism and pharmacodynamics.* Biochem Pharmacol, 2018. **153**: p. 196-204.