UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Characterisation of inducible proteins in the HepG2 cell line proteome

by

Andrea Antonio Ellero

A thesis submitted in fulfilment of the requirements for the degree

Doctorate of Philosophy

in

Pharmacology

Department of Pharmacology
School of Medicine
Faculty of Health Sciences
University of Pretoria

Supervisor: Professor AD Cromarty (University of Pretoria)
Co-supervisor: Doctor T Hurrell (Council for Scientific and Industrial Research)
Co-supervisor: Professor I van den Bout (University of Pretoria)

2022

Declaration

University of Pretoria

School of Medicine

Faculty of Health Sciences

Department of Pharmacology

I, Andrea Antonio Ellero

Student number: 13209656

Subject of work: Thesis

"Characterisation of inducible proteins in the HepG2 cell line proteome"

Declaration

1. I understand what plagiarism is and am aware of the University's policy in this regard
2. I declare that this thesis is my own original work. Where other people's work has been used, this has been properly acknowledged and referenced in accordance with departmental requirements
3. I have not used work previously produced by another student or any other person to hand in as my own
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work

Signature

.....

Acknowledgements

I would like to acknowledge the following people and entities for their contributions and support throughout the duration of this study.

I. The first of my thanks must go to my panel of supervisors, you have all been pivotal in my success and I would not be the ambitious researcher-in-training that I am today without your cumulative contributions to my development
   - To Dr Tracey Hurrell, thank you for facilitating my first exposures to research, for fostering my curiosity, teaching me how to think critically and guiding my endeavours. You have been an ever-present pillar in my research career, a brilliant mentor and friend. I hope to one day be able to do for others as you have done for me.
   - To Professor Duncan Cromarty, for constantly putting the success of all his students first. You have been a complex mixture of academic advisor, career guidance counsellor, motivator, psychologist and croc-wearing fiend. I truly appreciate all that you have done for me throughout the years and will strive towards becoming a researcher that you can be proud of.
   - I am grateful to Professor Iman van den Bout. Thank you for taking me in as a stray while Tracey and Duncan were abroad during building renovations and providing me the opportunity to conduct this research at the Centre for Neuroendocrinology where I was able to harness skills in all things cell culture. For being a constant source of scientific guidance and someone who has constructively challenged my ideas.

II. To my parents, Antonella and Nicola Ellero. Who have sacrificed a great deal to support me throughout my life and studies. I am sincerely grateful for all that you have done.

III. To Drs Ireshyn Govender and Previn Naicker, thank you for supporting this work through access to infrastructure and technical skills. You have gone from mentors to colleagues to dear friends and I am thankful of your contributions to my development.

IV. The National research foundation and University of Pretoria for financial support contributed towards this project.

V. To Drs Feng Wang and Christopher Walker who unknowingly mentored my coding skills and bioinformatics abilities

VI. Dr Mare Vlok who has been instrumental in the generation of high-quality proteomics data.

VII. To my friends and colleagues at the University of Pretoria spread across the Departments of Pharmacology, CNE, and ICMM it has been a great pleasure to work and study alongside you.

VIII. To my friends and family at large, specifically Hafiza Parkar, Shamiso Mlambo, Tarryn (Susan) Radomski, Adele Nel, Mandy Naude, Anastasios Athanasiadis, Konrad De Vries and Ane Pieters who have encouraged, supported and at times kept me sane throughout this journey.

IX. To Dr Tamara Romanini, my fiancé and soon to be wife. For your unwavering support, love, patience, and kindness throughout this journey.

X. Lastly, to my Nonni. Zio Achille, Zia Gabriella, my Nonna Olga, Nonna Matilde and Nonno Vincenzo who have all raised me to be the person that I am today. I dedicate this work to you.

"The only true wisdom is in knowing you know nothing." - Socrates

Table of Contents

Contents

Table of Figures

List of Tables

List of Abbreviations, Acronyms and Symbols

A

ACN	Acetonitrile
ADH	Alcohol dehydrogenase
AFP	α-fetoprotein
ALB	Albumin
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ACAP	Acetaminophen
APOB	Apolipoprotein B
ASGPR	Asialoglycoprotein receptor
AST	Aspartate aminotransferase
ATCC	American Tissue Culture Collection
ATP	Adenosine triphosphate

B

BCA	Bicinchoninic acid

C

C18	Silica-based column
Ca^{2+}	Calcium
CK18	Cytokeratin 18
CID	Collision-induced dissociation
CO_2	Carbon dioxide
$Cu^{2+/+}$	Copper
CXCR	CXC chemokine receptor
CYP	Cytochrome P450

D

D	Day
DAPI	4',6-diamidino-2-phenylindole
ddH_20	Double distilled water
DEPs	Differentially expressed proteins
DMEM	Dulbecco's Modified Minimum Essential Medium
dH_2O	Distilled water
DILI	Drug-induced liver injury
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol

E

ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization

F

FBS	Fetal bovine serum
FDA	Federal drug administration
FDR	False discovery rate
FGF	Fibroblast growth factor

G

g	Relative centrifugal force
GATA	GATA (coding region) binding protein
GSH	Reduced glutathione
GST	Glutathione S transferase

H

HBSS	Hank's balances salt solution
HCD	Higher energy collisional dissociation
HEPACAM	Human epithelial cell adhesion molecule
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HepG2	Human hepatoma cells
HGF	Hepatocyte growth factor
HLCs	Hepatocyte-like cells
hiPSC	Human induced pluripotent stem cells
HNF	Hepatocyte nuclear factors
HPLC	High-performance Liquid Chromatography
HSA	Human serum albumin
HSP	Heat shock protein
HTS	High-throughput screening

I

IAA	Iodoacetamide
ICP-MS	Inductively coupled plasma mass spectrometer
IEF	Isoelectric focussing
IGF	Insulin-like growth factor

IgG	Immunoglobulin G
IHC	International Conference on Harmonisation
INDs	Investigational new drugs
iPSC	Induced pluripotent stem cells
IS	Internal standard
iTRAQ	Isobaric tags for relative and absolute quantitation

L

LC	Liquid chromatography
LC-MS	Liquid chromatography–mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry

M

M	Molar (moles/liter)
MAPK	Mitogen-activated protein kinases
Mg^{2+}	Magnesium
$MgCl_2$	Magnesium chloride
MGF	Mascot generic format
miR	micro-RNA
mRNA	Messenger ribonucleic acid
MRP	Multidrug resistance-associated protein
MS	Mass spectra or mass spectrometric
MS1	Precursor ion scans
MS2	Product ions scans

N

NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
Na_2CO_3	Sodium carbonate
$Na_2C_4H_4O_6$	Sodium tartrate
NCEs	New chemical entities
NES	Normalised Enrichment Score
NSAID	Non-steroidal anti-inflammatory

O

OATs	Organic anion transporters
OCT-4	Octamer-binding transcription factor 4

P

PBS	Phosphate buffered saline

PCA	Principal component analysis
PDA	Photodiode array
pH	Power of hydrogen (negative logarithm of hydrogen ion concentration)
PHH	Primary human hepatocytes
PMSF	Phenylmethanesulfonyl fluoride
PTM	Post-translational modification

Q

QC	Quality control
qPCR	Quantitative polymerase chain reaction

R

R&D	Research and Development
rBM	Reconstituted basement membrane
RIPA	Radioimmunoprecipitation buffer
RLU	Relative luminescence units
RNA	Ribonucleic acid
RNase	Ribonuclease
RNA-Seq	RNA sequencing
ROCK	rho-associated protein kinase (Y-27632)
ROS	Reactive oxygen species
RP-HPLC	Reverse phase high performance liquid chromatography
RPLC	Reverse-phase liquid chromatography
RSA	Republic of South Africa

S

SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SERPINA	Serpin peptidase inhibitor, clade A
SILAC	Stable isotope labelling by amino acids in cell culture
SRM	Selected reaction monitoring
SRB	Sulforhodamine B
STs	Sulfotransferases
SYBR	SYBR Green I (DNA binding dye)

T

TCA	Trichloroacetic acid
TEA	Tris-acetate-EDTA buffer
TFA	Trifluoroacetic acid

TGF-β	Transforming growth factor-β
TMT	Tandem mass tags/tagging
Tris	Tris(hydroxymethyl)aminomethane
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride

U

UGT	Uridine 5'-diphospho-glucuronosyltransferase
USA	United States of America

Z

ZFNs	Zinc finger nucleases

Misc

2D	Two-dimensional
3D	Three-dimensional
Å	Angstrom
°C	Degrees celsius
%	Percentage
ΔΔCT	Delta Delta Ct
cm	Centimetres
Da	Dalton
eV	Collision energy
kDa	Kilo-dalton
mg/ml	Milligrams per millilitre
ml	Millilitre
ml/minute	Millilitre per minute
mm	Millimetre
mM	Millimolar
mmu	Milli mass units
ms	Milliseconds
m/v	Mass per volume (grams per 100 millilitre)
m/z	Mass per charge
µ	Micro
µl	Microliter
µl/min	Microliter per minute
µm	Micrometre
µmole	Micromole
µM	Micromolar
n	Nano
ηg	Nanograms

ηm	Nanometre
ppm	Parts per million
s	Seconds
v/v	Volume per volume (millilitre per 100 millilitre)

Abstract

Rationale: Attaining scientifically robust, clinically representative, and functionally translatable preclinical models for pharmaceutical research has become increasingly desirable. Relevant hepatic models are pivotal for the screening and accurate identification of safe new drug candidates. Currently, primary human hepatocytes are considered the gold standard for hepatic modelling However, their limited accessibility, high variability and terminal nature hamper their widespread adoption. Immortalised cell cultures, for example the human hepatoma cell line (HepG2), have served as widespread substitutes. However, the representativeness of HepG2 cells is of concern due to the relatively large phenotypic disparity when compared to primary human hepatocytes. Recent studies have postulated that the adoption of HepG2 cells into three dimensional (3D) conformations may promote a more closely correlated phenotype. Little is known regarding the extent of proteomic changes which occur when cells as cultured as spheroids, nor is there sufficient insight into the temporal dynamics which may drive these changes. The aim of this study was to investigate whether proteomic changes occur in HepG2 cells when cultured as spheroids compared to traditional monolayer culture techniques and if so whether these changes were culture time dependent. Additionally, HepG2 spheroid cultures were assessed for improved metabolic adaption following extended-time spheroid-based culture in the presence of specific drugs, a key characteristic of human hepatocytes that would be critical for hepatotoxicity testing models.

Methods: HepG2 cells were cultured as spheroids using two established 3D culture models; cells were seeded into hanging drop plates (20 000 cells/spheroid) and into 81 well, 3D micro moulds (1000 cells/spheroid), then maintained under untreated control or drug exposed groups, using a 7-drug cocktail representing substrates of phase I metabolizing enzymes. Spheroid cultures were characterized for protein content, size, viability, and immunohistochemistry. Cultures from different groups

and culture times were compared by means of quantitative proteomic methods using isobaric labelling proteomics. Sample proteins were reduced, alkylated, precipitated, digested with trypsin and labelled using 6-plex tandem mass tag (TMT) isobaric labels. Proteomic analysis was performed using a Dionex Ultimate 3000 RSLCnano liquid chromatography system coupled to a Thermo Scientific Fusion Lumos Orbitrap Mass Spectrometer. Protein identification was conducted using an alternating SequestHT/MS Amanda schema. Relative protein quantitation for the different spheroid groups were expressed relative to original monolayer cultured HepG2 control cells. Relative metabolic competences of spheroid cultures were assessed using a defined hepatic metabolism phenotyping cocktail using an in-house developed liquid chromatography tandem mass spectrometry method.

Results: HepG2 spheroid cultures were shown to progressively upregulate expression of hepatic marker proteins over the 28-day culture time course. Quantitative proteomic methods initially identified over 5000 proteins which when filtered to include only proteins with at least three identified peptides and presence in all replicates, reduced to approximately 4800 proteins. Identified proteins showed differential expression between spheroid culture groups compared to monolayer cultures. Additionally, progressive differential proteome expression was evident between spheroid cultures according to time spent in culture and based on presence or absence of drug exposure. Thorough bioinformatics analyses revealed insights into the potential mechanisms responsible for the proteomic differences and identified that while spheroids exposed to drug appear to adapt metabolically to their culture environments, the nature of these changes may lack impact for clinical relevance. Metabolomic investigations corroborated the proteomic findings showing that, while maintaining some capacity for phase I metabolism, overall, clinically relevant metabolism for cytochrome P450 enzymes was absent.

Conclusion: Quantitation of large protein cohorts demonstrated that there are adaptive changes occurring during long-term culture of HepG2 cell spheroids. While

the proteomic changes observed, under all culture conditions, do not support adoption of these cultures for modelling drug metabolism, use in other bespoke model systems which concern other aspects of liver function is feasible. The high-quality proteomic data generated shows temporal dynamics in spheroid cultures which have implications for their utility. It necessitates the requirement for more comprehensive investigations into the relevance of the time-points used when adopting spheroid cultures.

Chapter 1. Literature review

1.1. The drug development process

The rise and distribution of modern medicine has led to a global increase in average life expectancy for humanity [1]. This increase is attributable to the continuous emergence of new drugs and therapeutics which facilitate the improvement of patient prognosis and in the best cases, elimination of disease. A drug can be defined as a substance which produces a physiological or psychological effect when introduced to an organism. It is useful for its diagnosis, treatment or management of a disease or pathological condition [2]. The drug development process is a time consuming and high-risk process. The development of a single new therapeutic often carries costs which surpass two billion USD, and usually require the associated financial burden to be borne by both the private and public sectors [3, 4]. The average time taken for the discovery, development and approval of new drug candidates is approximately 13.5 years. For approval to be granted, the drug candidate must first demonstrate both safety and efficacy [5]. A general strategy adopted by pharmaceutical companies is to initially invest in research to understand disease development, progression and pathology in order to identify therapeutic targets for potential treatments [5]. The drug development process is made up of 4 phases namely, discovery, preclinical research, clinical research, and review by drug regulating authorities. The process begins with the discovery phase which involves review of the existing data surrounding specific disease conditions and potential drug targets, then screening large compound libraries of up to 100 000 compounds for potential candidate drugs. These screenings are often initially *in silico*-based molecule docking studies [6, 7]. The discovery phase is followed by the preclinical phase, whereby the promising compounds identified during screening are evaluated using *in vitro* modelling and testing followed by *in vivo* animal studies. Discovery and preclinical testing are routinely conducted concurrently and may take up to 7 years to complete.

Compounds showing promising efficacy and low toxicity are selected, and an application is made to the drug regulating authorities for permissions to undertake a clinical trial. Phase I clinical trials begin, to address drug safety, in healthy volunteers and typically include only small numbers of participants (20-50). Phase II trials, which can be randomized, evaluate the candidate drug's efficacy and establish a side effects profile over a period of 1-2 years in 40-120 participants. Finally, development of the unapproved candidate drug is concluded with phase III clinical trials, that typically last 3-4 years and the number of participants range from several hundred up to many thousands. The purpose of phase III trials is to confirm treatment efficacy and to monitor for long term side effects and potential drug interactions in a diverse population. Phase IV, or post-approval surveillance, should take place after the drug regulating authority gives marketing or emergency use approval to the drug candidate. [8].

Clinical trials are encumbered by strict regulatory and ethical requirements yet show low success rates and successful drug development remains a slow and costly process. In general, over half of all drugs fail during phase II and phase III clinical trials due to a lack of efficacy whereby the drugs are unable to demonstrate a clinical benefit, and about another third of drugs fail due to safety issues including a risky narrow therapeutic index where small discrepancies from the intended drug exposure may lead to poor efficacy or toxic side effects [9]. Hwang *et al.* reviewed 640 phase III trials with novel therapeutic agents and found that 54% failed during clinical testing with 57% of those being due to inadequate efficacy [10], while approximately 17% failed due to safety concerns. Though safety is monitored through every phase of clinical trials, potential complications generally only become apparent during the larger study population sizes included in phase III trials or, more detrimentally, during post-marketing surveillance (phase IV) [11]. By this point, a large proportion of available resources would have been committed. An overview of the drug development process is summarized in Figure 1.1 that indicates the phases and the attrition rate of potential candidate drug compounds.

Drug safety is not always straightforward to assess, as inter-patient genotypic and phenotypic variability are major contributors to drug exposure outcomes and influence the progression of certain diseases or conditions and may be dose or time since exposure dependent. Genotypic considerations are important to ensure the safety of participants in a clinical trial, however, a duality exists in their application. Screening of participants for potential risk factors and pre-trial exclusion may aid in the preservation of both resources and to reduce overall participant risk.

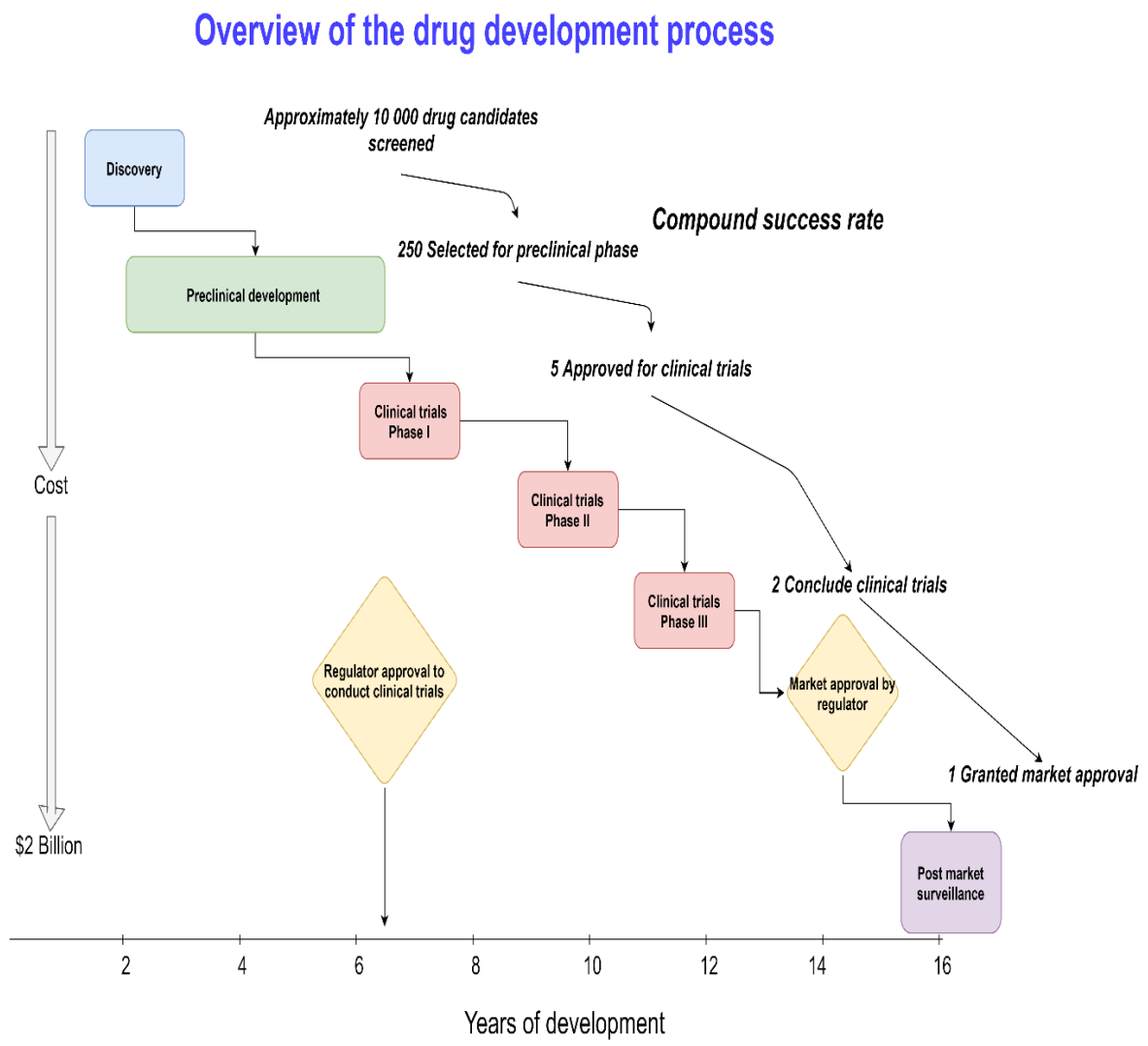


Figure 1.1 Overview of the drug development process. Figure produced using Diagrams.net.

An example is described in a personal communication by Fijal *et al.* [12] where, the genotypic screening of over 600 participants for three genetic "poor metabolizer"

phenotype variants of CYP2D6 prompted the exclusion of 40 individuals from a trial thereby reducing the sponsor cost by approximately $300,000 while simultaneously reducing the overall safety risks to the participants in the study. Conversely, improper disclosure of these approaches may be used to bias participant safety data which inappropriately represents a drug under investigation as safer than it truly is. The use of these screening methods, while potentially useful to the development pipeline, do not entirely circumvent the potential for wasted investment of a trial therapeutic which will ultimately fail as these interventions are only possible during the controlled later phases of drug development. The earlier in the pipeline a drug under investigation fails, the lower the wasted development costs will be.

1.2. The role of preclinical screening

The Nuremberg Code and Helsinki Declaration dictate minimum research ethics surrounding human experimentation, which mandates that any new experimentation to be conducted on humans be based on prior data from animal studies [13]. From this mandate, the preclinical phases of drug development were established and are now an international legal requirement. Ideally, investigational drugs would be screened for safety and efficacy earlier during the development process, using *in vitro* and preclinical animal experiments where intact neural, immune, circulatory, hepatic, gastric, renal, and excretory systems can be affected. However, this screening effort is confounded by the lack of representative models with adequate predictive potential for accurate toxicity or efficacy testing. It is therefore essential to improve on the accurate predictability aspects of *in vitro* type models. Modelling for preclinical studies initially involves *in vitro* analysis in selected representative cellular models, followed by non-clinical *in vivo* studies in three different species of relevant animal models. The purpose of these platforms is to establish metabolic and toxicity profiles and to determine pharmacokinetic and pharmacodynamic parameters. Pharmacokinetic parameters investigated must include absorption, distribution,

metabolism, and excretion (ADME) characteristics which are essential in understanding safety concerns and determining drug concentrations that can be expected at the target site that influences the efficacy. Historically, the majority of high throughput screening (HTS) for drug safety and efficacy have taken place using two-dimensional (2D or monolayer) cell culture model systems [14]. While useful in their ease of applicability, relatively low cost, and high turnover, 2D systems suffer several notable limitations when addressing the requirements of reliable model systems [15]. These limitations include rapid loss of cell polarity, limited cell-cell and cell-extracellular matrix (ECM) signalling [16]. Cells cultured in 2D undergo comparatively rapid growth compared to their *in vivo* counterparts, where disparity exists with regards to oxygen gradients, utilization of nutrients and growth factors as well as build-up of metabolic waste products [17, 18]. Several cell types undergo rapid de-differentiation and experience loss of function when cultured in 2D, as in the case of primary human hepatocytes (PHH) [19, 20]. These characteristics have prompted a paradigm shift within recent years where 2D systems are being substituted with three-dimensional (3D) spheroid systems [19-22]. Three-dimensional culture models allow cells to better recapitulate their *in vivo* architecture. By their nature, 3D systems allow a higher proportion of cells in direct contact with other cells and promotes the maintenance of cellular polarity as well as deposition of ECM [16, 23]. Three-dimensional systems still carry several limitations; the formation of spheroids using single cell types, an absent vascular system and supporting structures which underrepresent the inherent physiological complexity. Prior to clinical investigations in humans, animal models with fully intact physiological systems are used to improve relevance. These platforms will be discussed more extensively later in this chapter.

1.3. Contextualising in vitro and preclinical modelling platforms

Revisiting one of the initial steps in drug candidate development, *in vitro* cultures are established on the premise of simplification. *In vitro* models have been utilized for

over 100 years and were originally described by Loeb in 1897 when maintaining rudimentary cultures of liver, kidney, thyroid and ovary *in vitro* on plasma clots for up to 3 days [24]. A variety of commercially available healthy or diseased tissue cells have since been adapted and optimised for use in culture and are typically grown in 2D formats, generally on sterile polystyrene-based culture vessels that can be also be sourced commercially [25]. Simple 2D tissue cultures provide a highly controlled and easily reproducible growth environment but sacrifice their *in vivo* mimicry in favour of ease of experimental facilitation. Despite this, these systems are valuable for measuring initial parameters related to efficacy or toxicity. These include elucidating molecular mechanisms of action and investigation of the ligand-target interactions of drugs as well as potentially identifying off target effects [25]. Many aspects of cell biology are observable within homogenous cell type 2D *in vitro* cell cultures and are robust despite the absence of the relevant supporting physiological systems.

Primary cell cultures are directly derived from fresh tissues or organs, and in many applications are regarded as the gold standard as they most accurately mirror the biology of the tissues from which they originated. However, primary cells are limited in their availability, are expensive, difficult or slower to expand and in many cases cannot be expanded or passaged more than 3 – 4 times and can be highly variable depending on the source and method of harvesting. They are also often contaminated by supporting stromal cells and exhibit varying levels of dedifferentiation within hours or days when cultured in 2D. Finally, primary cells are more susceptible to passaging-induced phenotype alterations, are increasingly sensitive to their culture environment post-harvest and typically require more sophisticated techniques or complex culture media composition.

Primary cell organoids are 3D conformations of mixed primary cell types derived directly from a specific organ or tissue type [26]. Sources are typically explanted tissue harvested during surgery or from patient biopsies. These cultures retain the histological architecture and genomic background relevant to each donor cell type in the organoid and have been shown to be more representative of the heterogenous

nature of patient tissue [27, 28]. The growth and maintenance of organoid cultures requires cell specific growth factor cocktails with varying combinations of Wnt, R-Spondin-1 (a Wnt amplifier), epidermal growth factor (EGF), prostaglandin E2, fibroblast growth factor 10 (FGF10), noggin (inhibitor of bone morphogenetic protein signalling), A83-01 (inhibitor of TGF-signalling), SB202190 (p38 inhibitor) and Y-27632 (Rho/ROCK kinase inhibitor) with new, more tailored cocktails constantly being optimised [29]. Relatively slow growth rates, patient specific variability, labour intensiveness, requirement of highly skilled personnel, limited cell sources, and difficulty in upscaling to HTS are present challenges in the mainstream adoption of primary cell organoids [30].

Transformed (immortalized) cell lines were established over 30 years ago and are still commonly utilized in preclinical settings. Transformed cells are originally derived from primary cells but have been genetically altered to overcome several culture related challenges associated with primary cells. Transformation is achieved by the introduction of oncogenes into the cells genome to confers characteristics such as faster proliferation rate, increased resistance to external stressors and a reduced rate of dedifferentiation [31, 32]. Indeed**,** these adaptations carry several knock-on limitations; increased growth kinetics impose an increased demand on cellular resources which often causes a cell to switch from a functional phenotype to a proliferative phenotype. Though phenotypically stable in the short term, genotypic drift is a common occurrence within long term passaged cells which leads to high inter-lot variability and poor reproducibility between older and newer cell batches [33, 34]. Transformed cell lines are robust and therefore more successful in transfection studies, allowing the stable introduction of genes in instances requiring the expression of a specific reporter, inducible expression of an extrinsic protein or as an overexpression system. Conversely specific methods of gene suppression (antisense RNA, small interfering RNA) allow the inhibition of gene activity more easily [35].

The concept of differentiating one cell type into another cell type was pioneered by developmental biologist Sir John Gurdon who attempted the first nuclear transplantation in 1958 [36]. Since then, the field of developmental biology has undergone extensive advances including the establishment of induced pluripotent stem cell (iPSC) cultures. Takahashi and Yamanaka identified several transcription factors (OCT-3/4, SOX-2, cMYC, and KLF-4) involved in the pluripotency of cells [37]. A multitude of differentiation protocols have since been established, enabling the differentiation of iPSC's into numerous cell lineages/types such as neuronal, hepatic, gastrointestinal etc. [38, 39]. These cultures have been used in an array of applications including organ and disease modelling, toxicology, embryology, and developmental biology. However, inconsistent differentiation endpoints, emergence of alternate cell types during differentiation, potential for divergence to a cancer-like phenotype and the often-attained foetal characteristics associated with iPSC derived cultures remain limitations which confound their biology.

Animal models have been serving as the front-line testing platforms to simulate human physiology and modelling an *in vivo* biological response to drugs in development. Data from animal studies confirming safety of drugs are used as the premise for initiation of first in human (FIH) phase 1 clinical trials [40]. Animal models are often used to demonstrate modulation of disease progression and to develop rudimentary proof-of-concept efficacy and baseline safety [41]. Although animal models are an essential platform for establishing these parameters and mitigating harm to humans during drug development, the general concordance between humans and animals remains low. Altered volume of distribution, variable basal metabolic rates, discordance in receptor expression, and differential metabolic pathways all serve as limitations to animal models. More recently, animal models have been genetically modified to express humanized gene products and to carry humanised organ systems [42]. Chimeric animal models are being developed to bridge some of the gaps presented in the use of standard animal models. However, these chimeras

are extraordinarily expensive to produce and carry considerably more ethical concerns depending on the extent of their humanization.

In general, when considering the current states of preclinical development platforms, it is apparent that no single model system exists which is capable of modelling all facets required for the development of new drugs. Careful consideration is required when selecting the most appropriate model system or combination of systems. The understanding of what data/endpoints a particular model can accurately and relevantly provide are essential to addressing the challenges arising during drug development, as well as deciding on the trade-offs between HTS adaptability and physiological relevance. Figure 1.2 shows the increasing complexity of available cell culture and preclinical animal models with increasing relevance to the human contrasted against the ability to perform HTS. Table 1.1 provides a brief overview of the of advantages and limitations associated with the various model systems.

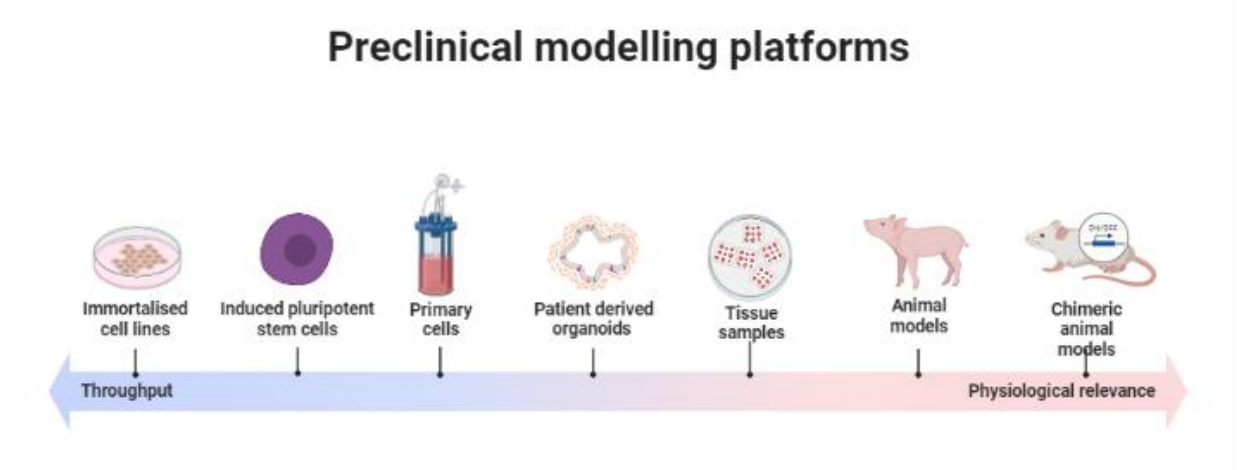


Figure 1.2: Preclinical modelling platforms comparing physiological relevance with adaptability to high throughput analysis. Created with BioRender.com

Table 1.1: Summary of advantages and disadvantages associated with preclinical models

Model	Advantages	Disadvantages
Immortalized cell lines	• Simple and low-cost maintenance • Long term culture • Amenable to genetic manipulation • Mainly valuable for drug screening and gene expression patterns	• Lack of relevance compared to *in vivo* microenvironment • Growth kinetics not representative • Loss of cell polarity • Loss of correlation with host tissue • Cancerous phenotypes
Primary cells	• Replicate patient specific phenotypes • Applicable *in vivo* functional correlation	• Limited culture duration • Limited supply • Inter-donor variability • High cost
Induced pluripotent stem cells	• Multiple somatic cells can act as cell sources • Donor genetics preserved • High proliferative/ differentiation capacity • Useful in understanding developmental pathways • Highest promise for preclinical modelling	• Differentiation protocols not fully refined • Immature differentiation phenotype • Inconsistent terminal phenotypes • High cost • Requires specialized personnel • Prone to differentiating heterogenous cell types • Ethical concerns in the case of derived models for developmental biology applications • Laborious culture protocols
Patient-derived organoids	• Highly representative of patient phenotypes • Ability to mimic organ ultrastructure • Suitable for precision medicine • Can be cryopreserved for bio-banking • Relatively genetically stable	• High cost • Limited cell sources • Inter-donor variation • Inter-culture structural variation • Less amenability to HTS • Requires specialized personnel • Laborious
Animal models	• Full organism system • Suitable for ADMET studies • Ability to track multiple organ responses simultaneously • Multiple routes of drug exposure • Useful for studying development/embryology	• Lack of relevance to humans • Ethical considerations • Laborious • Can be costly • Require housing facility and trained personnel
Chimeric animal models	• Similar to animal models • Higher degree of correlation to humans • Ability to modulate human pathologies	• Similar to animal models • Extremely costly • Difficult to adapt to HTS

1.4. Safety pharmacology and the liver

Drug related toxicity is a common occurrence in the pharmaceutical industry. Improper dosing, patient deviation from prescribed regimens, genetic variance, and drug-drug or drug-food interactions are all factors that may contribute individually or concurrently to a toxic response, even where drugs have reported wide margins of safety. The concept of dose related toxicity was pioneered over 500 years ago by Paracelsus [43], with the notion that all substances are poisonous, but that the dose distinguishes a poison from a remedy. A 'margin of safety' for drug doses must therefore be established to prevent toxicity. Presently, preclinical animal testing is needed to establish the no-observed-adverse-effects-level (NOAEL), which is the highest dose at which no signs of toxicity are detected in an animal model [44].

Toxicity can be divided into at least four categories according to the pathological effect induced. These include direct cell death/tissue injury, altered phenotype/function, immunological hypersensitivity, and mutagenicity/carcinogenicity. In a similar way, drug mediated toxicity can be divided into five categories, namely, mechanistic or target related toxicities, immunological hypersensitivities, off-target site or peripheral toxicities, idiosyncratic toxicities, and biological transformation induced toxicities [44].

Over 90% of adverse clinical reactions are associated with frequent use of drugs such as analgesics, anticoagulants, anticancer drugs, antimicrobials, antidiabetics, diuretics, and steroids. Most cases of adverse clinical reactions manifest as exaggerated effects at the primary drug target (mechanism-based toxicity) due to; dosing errors, prolonged and/or permanent use, and cytotoxicity. Hepatotoxicity is one of the most notable toxic effects since the liver is responsible for metabolizing most exogenous compounds including most drugs [45]. This toxicity may go unnoticed for extended periods of time due to the size of the liver and the ability to repair itself under mildly toxic conditions. Many drugs with hepatotoxic effects may only exhibit obvious adverse effects long after the drug was first administered or even after discontinuing the drug.

There are many mechanisms which may constitute or contribute to hepatotoxicity [46]. The underlying mechanisms initiating hepatotoxicity influences the extent of liver damage and alters specific markers commonly used for identification/diagnosis of the extent of hepatic damage. These markers are also used to monitor treatment success. Reactive drugs or drug metabolites produced as intermediates during drug metabolism may bind covalently to critical intracellular proteins leading to cellular dysfunction. This causes cytoskeletal disruption and a decline in adenosine triphosphate (ATP) that results in swelling and cell membrane rupture (Figure 1.3A) [47]. Drugs that block transport proteins in hepatic canaliculi can interfere with bile and metabolite export (Figure 1.3B) resulting in cholestasis, which when combined with other underlying pathologies such as pre-existing genetic defects of supporting proteins may cause a secondary hepatotoxicity [48]. Biotransformation of drugs can involve reactions where drug-enzyme adduct formation occurs spontaneously, decreasing the available metabolic capacity as these adducts are functionally inactive. These adducts may be moved to the hepatocyte cell surface where they serve as immunogens, stimulating the production of antibodies or inducing direct cytotoxic T-cell responses (Figure 1.3C - D) [49]. The resulting immune cell cytokine response can elicit programmed cell death (apoptosis) driven by ligand binding to tumour necrosis factor (TNF) and CD95 (Fas) receptors (Figure 1.3E) [50]. Drugs can also bind to and disrupt mitochondrial DNA or hinder the cellular respiratory pathways or disrupt fatty acid oxidation blocking energy production. This can cause steatosis type liver injury (Figure 1.3F) [51].

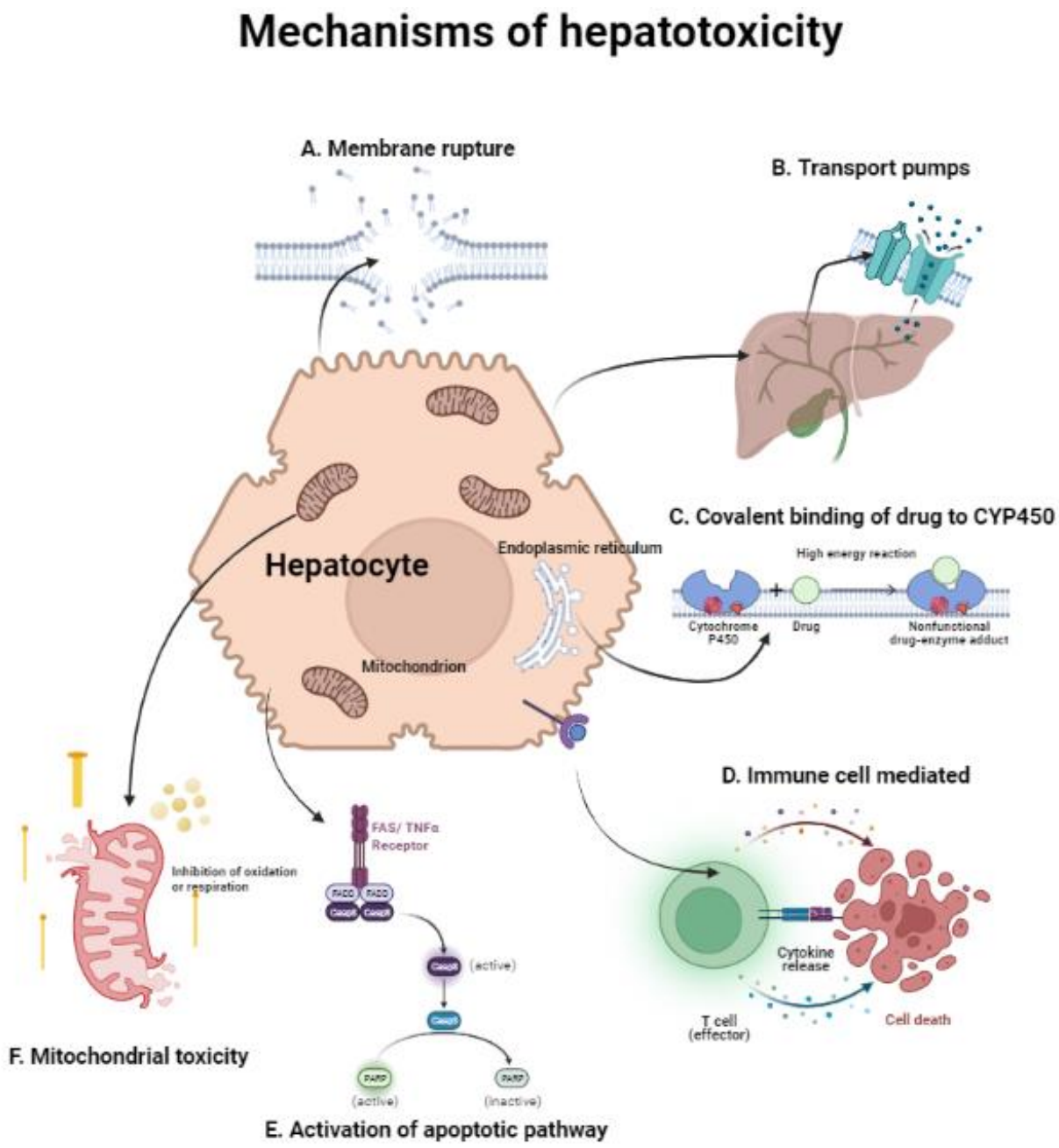


Figure 1.3: Mechanisms driving hepatotoxicity. Created with BioRender.com

1.5. The liver as an adaptable organ

The liver is a compact multilobed granular organ with a high regenerative capacity. Hepatocytes are parenchymal cells which together with biliary epithelial cells (cholangiocytes) constitute approximately two thirds of the hepatic cell population. These cells are spatially heterogenous along the portal-central axis of the liver lobule,

with zonation being evident ultrastructure level [52, 53] at the with enzyme activities and other cellular functionality grading across the zones of liver lobules. Non-parenchymal cells include hepatic sinusoid epithelial cells, hepatic stellate cells, vascular cells and Kupffer cells (a subpopulation of differentiated hepato-specific macrophages). During development, post-gastrulation, the major hepatic cell types (hepatocytes and cholangiocytes) are derived from the endoderm with the remaining cell types deriving from the mesoderm. Foetal liver cells known as hepatoblasts form the liver bud which follows a series of migratory and proliferative processes giving rise to the liver organ. During the epithelial differentiation stage these hepatoblasts either mature into hepatocytes or differentiate into cholangiocytes [54].

In the mature non-foetal liver, the hepatoblast cell population no longer exists, however the liver maintains the capacity for regeneration, which is a characteristic absent in most mature organ types. This regenerative capacity is enabled by a plasticity unique to liver cell populations [54-57]. Historically it has been assumed that hepatic regeneration is driven by an unknown underlying population of dormant hepatic progenitor cells. However, more recently, multiple lines of evidence suggest the main mechanism of hepatic regeneration is via the dedifferentiation of hepatic cells, which were assumed to be terminally differentiated (eg. hepatocytes, cholangiocytes), into cells with progenitor-like functionality which then form the bulk of the replenishing cell source. Figure 1.4 represents the proposed model of cellular differentiation, dedifferentiation and transdifferentiation of hepatocytes, cholangiocytes and liver progenitor-like cells to facilite cellular replenishment during liver injury.

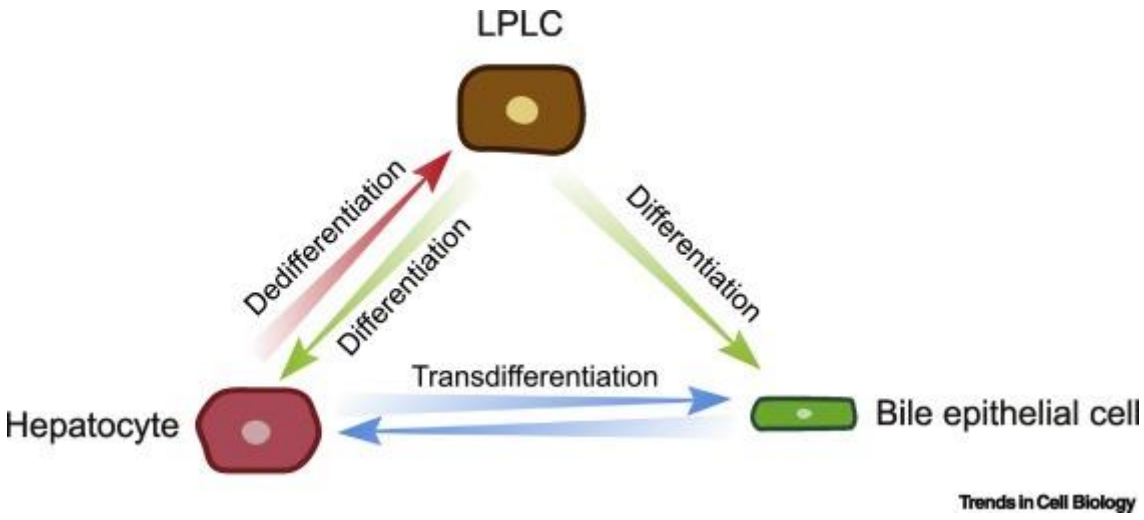


Figure 1.4: During liver injury, hepatocytes undergo dedifferentiation to liver progenitor-like cells (LPLCs), which express both hepatocyte and biliary epithelial cell markers (red arrow). LPLCs differentiate into hepatocytes and biliary epithelial cells to regenerate the liver after injury cessation (green arrows). In mammals, when liver injury becomes severe, biliary epithelial cells can transdifferentiate into hepatocytes via a biphenotypic status expressing both biliary epithelial cell and hepatocyte markers (blue arrow). Similarly, hepatocytes can transdifferentiate into biliary epithelial cells to regenerate biliary epithelium when these cells become depleted (blue arrow) reprinted from [58] with permission from Elsevier under the Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND).

This plasticity is likely to be governed by a combination of intrinsic and extrinsic cell mechanisms such as the retention of progenitor permissive chromatin by hepatocytes conferring a degree of reprogramming competence [59]. The cellular microenvironment and peripheral cell signalling may contribute to the plasticity of mature hepatocyte populations during certain subtypes of liver injury.

1.6. Pharmacological relevance of the liver

Many tissues can metabolize drugs, but the most active tissue per unit weight is the liver [60]. The importance of liver metabolism is especially evident by its effective removal of lipophilic compounds from the body. This is achieved by metabolism to more polar and hydrophilic compounds which increases the rate of excretion by the kidneys. Without this metabolism, these compounds could easily accumulate in lipid rich tissues. The liver has an extraordinary, varied group of enzymes, which are

responsible for the metabolism of both endogenous and exogenous compounds. Drug metabolism is usually divided into two major sequential phases simply termed phase I and phase II reactions [60] which are preceded by phase 0 and followed by phase III reactions. The adequate expression and complex interplay of these coordinating enzymes is essential in metabolic functions.

Phase 0 reactions have been described as the initiatory transport reactions whereby the drug is transported from the blood into the liver and delivered to the hepatocytes through the basolateral uptake process [61].

Phase I reactions activate or prepare drugs for downline conjugation by adding polar groups to the structure of lipophilic drugs. This is achieved through reactions such as oxidation, reduction, hydrolysis, demethylation, and hydroxylation by the haem containing cytochrome P450 (CYP450) superfamily of enzymes. The CYP450 enzymes are integral membrane proteins primarily localized in clusters on the endoplasmic reticulum or in specialized vesicles known as microsomes [62] which serve to export hepatic enzymes to their functional zones. These enzymes are rapidly and highly inducible within the liver following exposure to a number of general or specific enzyme inducing drugs.

Phase II reactions are conjugation reactions whereby a polar molecule is covalently bound to the polar handle added during phase I metabolism. Glucuronic acid is the most commonly added group resulting in glucuronidation [62] but also includes sulphate or amino acid additions. These conjugations allow the whole molecule to become more hydrophilic which increases water solubility that enhances renal excretion.

Phase III reactions involve the transport of the metabolically transformed drug out of the cell [63]. These reactions do not change the structure of the parent compounds; however, these proteins are essential in the successful elimination of derivatized xenobiotics from the cell. The major family to which these transporters belong is the ATP binding cassette protein superfamily (ABC), which are found within cell membranes throughout all cells of the body. Notable transporters in the liver include

multidrug resistance associated protein 2 (MRP2) organic anion transporting polypeptide 2 (OATP2) and P-glycoprotein (Pgp or ABCB1) [64].

Cytochrome enzymes are considered the body's major group of drug metabolizing enzymes and were first identified by Martin Klingenberg in 1958. The human cytochrome P450 monooxygenase family is a membrane-associated, hydrophobic enzyme system with 57 reported members. These are contained in 18 families and 43 subfamilies. Families 1, 2 and 3 make up approximately half of the total number of CYPs in mammals. They are generally designated as drug detoxification enzymes due to their ability to metabolize many different xenobiotic compounds alongside the capacity to metabolize endogenous compounds. Members of Family 4 are better known for fatty acid ω- and (ω-1)-oxidation, rather than xenobiotic metabolism. Members of Families 5–51 (14 families) perform various physiological functions involving the metabolism of specific endobiotics [65].

The major CYP isoforms, which are responsible for approximately 75% of hepatic drug metabolism, in typical decreasing order hepatic expression are CYP3A4/5, CYP2D6, CYP2C8/9, CYP1A2, CYP2C19, and CYP2B6. Additional isoforms include those responsible for the metabolism of sterols (CYP1B1), vitamins (CYP2R2), eicosanoids (CYP4F2), fatty acids (CYP2J2) and others (CYP2A7) with currently unknown function [66]. In the liver, their prevalence is highly variable. Table 1.2 contains the major CYP enzyme families, their contribution to hepatic enzyme abundance, and the extent of their variability. Biological variance such as age, gender, race, disease state or enzyme induction are known contributors to the observed variability in these enzymes. [67].

Table 1.2: Representation of CYP enzyme proportion and extent of variability [68]

Human CYP Enzymes Reported to Be Involved in Metabolism of Xenobiotics Located in the Liver[a]

CYP enzyme	Level of the enzyme (% of total)	Extent of variability in level
1A2	**ca. 13**	**ca. 40-fold**
1B1	< 1	
2A6	ca. 4	ca. 30–>100-fold
2B6	ca. < 1	ca. 50-fold
2C	**ca. 18**	**25- to 100-fold** (depending on the enzyme)
2D6	**Up to 2.5**	**>1000-fold**
2E1	**Up to 7**	**ca. 20-fold**
2F1[b]		
2J2[b]		
3A4	**Up to 28**	**ca. 20-fold**
4A[b]		
4F[b]		

([a]Bold type indicates enzymes involved in the metabolism of numerous drugs. Underlining indicates enzymes involved in the metabolism of major chemicals with possible toxicological consequences. [b]No data.)

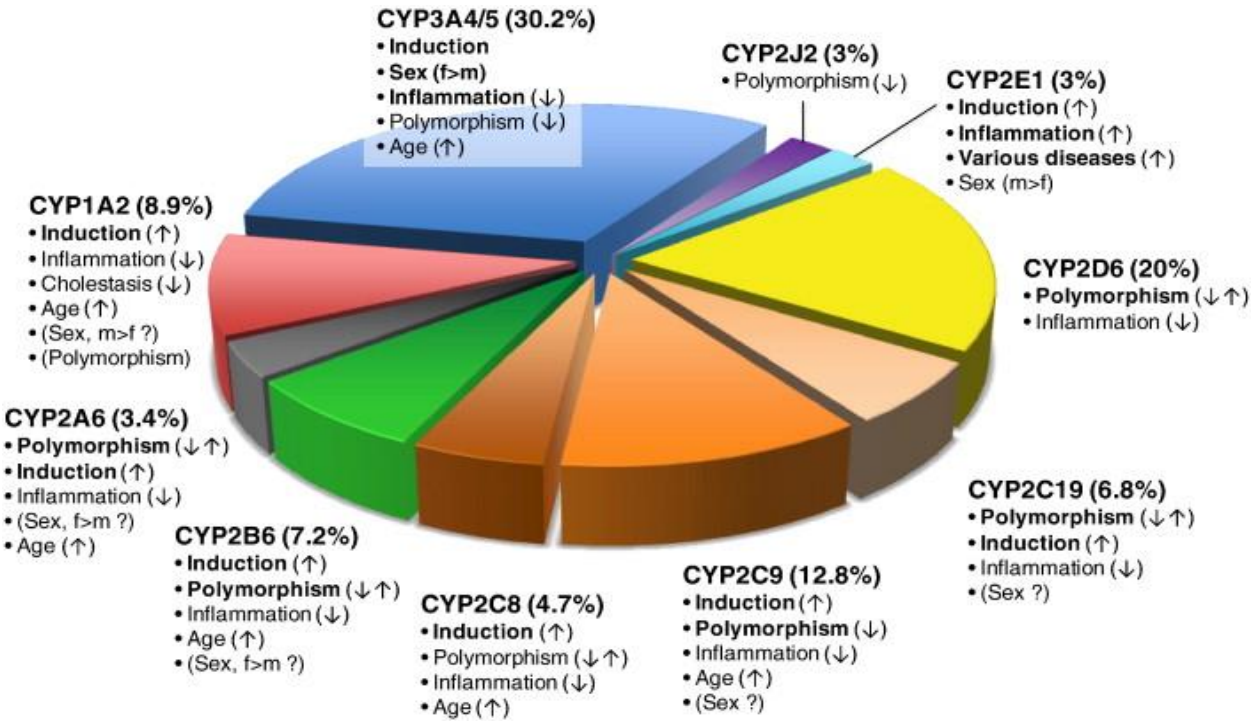


Figure 1.5: Proportion of commercial drugs metabolised by CYP isoforms, Adapted from [69] and reprinted with permission from Elsevier under the CC BY NC ND licence.

Although CYP450 enzymes exist in certain proportions in the liver, their role in drug metabolism is variable and does not correlate to their relative proportion. Figure 1.5 shows the percentage of commercial drugs metabolized by each major isoform. Regulation of drug metabolizing enzymes is carried out by an array of ligand-dependent nuclear receptors such as the aryl hydrocarbon receptor [70]. Upon activation, receptors are internalized and translocate to the nucleus where they then up-regulate the expression of phase I/II metabolizing enzymes [71] that alters the rate of specific drug metabolism pathways.

While metabolic functions of CYP450 enzymes are intended to be cytoprotective, there are instances where altered metabolism may lead to highly reactive products which can overload cell detoxification machinery and damage the cell [44]. Due to the dominant hepatic localization of these drug metabolising enzymes, potential liver pathologies are common. Under appropriate dosing regimens and absence of genetic variants, hepatotoxicity is rare, however, enzyme induction or inhibition or genetic variants can alter the prevalence of toxicity [44].

Hepatic expression of a single or multiple CYP families or subfamilies may be induced or inhibited following exposure to specific compounds. For instance, Midazolam primarily induces CYP3A4 whereas Phenobarbital is a broad-spectrum inducer which influences several different CYP families simultaneously. Mechanisms of induction may be variably attributed to several factors such as an increase in basal transcription or translation of enzyme mRNA, inhibition of enzyme degradation or increase in enzyme stability [72]. By making use of these properties, it is possible to increase the expression of certain enzyme subfamilies through selective induction with a drug cocktail.

1.7. In vitro hepatotoxicity screening methods

As attrition rates in drug development due to hepatotoxicity remain high, there is a need for new technologies that facilitate improved outcomes in drug discovery. Unlike other organ toxicity models, the concordance between human and animal

livers is low [23, 73] resulting in development of a number of *in vitro* models for screening for hepatotoxicity. Although failing to represent the complexity of a whole organism, *in vitro* models have been used extensively to investigate hepatotoxicity and have significantly contributed to the current understanding of hepatotoxic mechanisms [74-76].

Among *in vitro* systems, primary hepatocytes are recognized as the gold standard for toxicity screening due to their retention of most hepato-specific functions such as gluconeogenesis, glycogen metabolism, urea formation, plasma protein synthesis and secretion, lipid metabolism, and drug-metabolizing capacity (phase I and II enzymes) for limited periods in culture [77]. Consensus regarding cell culture conditions for hepatocytes used in toxicity screening is lacking despite this having a variable impact on hepatocyte function.

Standard culture conditions make use of chemically defined media with hormone supplementation and extracellular matrix coated culture surfaces, which are suitable for approximately one week of culture. Extended culture times require additional supplementation with soluble factors or the adoption of 3D culture conformations to maintain hepatic functions. [78]. Suspension cultures of hepatocytes have been widely used for monitoring toxicity of soluble compounds, however these cultures have only been useful in very short studies. Precision-cut liver slices preserve the intact liver structure and allow for simultaneous assessment of tissue morphology, biliary function, and response to potential toxicants. However, liver slices are confounded by the high inter-preparation variability and widespread scarcity which limits adoption of this technique [79]. Subcellular fractions have been used widely for rudimentary metabolic studies. However, in comparison to cellular models their hepatotoxicity screening capacity is limited [80].

Table 1.3 summarizes the advantages and limitations associated with several hepatotoxicity screening platforms using different primary human hepatocytes, subcelluar fractions and transformed hepatocytes.

Table 1.3: In vitro *hepatotoxicity screening platforms*

In Vitro Model	*Advantages*	*Limitations*
Liver slices	• Structure and metabolic complement maintained • Maintains cellular heterogeneity • Useful for histology • Functional assessments • Organ polarity preserved • Hepatobiliary function	• Short term viability • Scarce availability • High inter-donor and technical variability • Preservation and expansion limited • Difficult to culture
Primary hepatocytes (suspension cultures)	• Cryopreservation • Representative phenotype • Genotypically identical to donor • Suitable for functional studies (enzyme activity) • Inter-species assessments	• Short term viability (approx. 6 hours) • No cell-cell/cell-substrate interactions (may alter signalling pathways) • Unnatural culture environment • High variability • Limited accessibility for human isolates
Primary cultured hepatocytes (adherent cultures)	• Cryopreservation • Co-culture • Suitable for toxicology and function studies • Representative phenotype • Suitable for metabolomic studies • Cell-substrate and cell-cell (3D cultures) interactions • Inter-species assessments	• Phenotype instability in monolayers (less prevalent in 3D culture) • Limited cryopreservation cycles • High variability • Strong influence from culture environment on function/phenotype
Liver-derived cell lines (immortalized)	• High capacity for expansion • Easy to use • Cheap to maintain • Adaptable to HTS • Genetic manipulation • Suitable for pathway/receptor studies • High reproducibility	• Limited hepatic function • Low metabolic competence • Poor correlation to *in vivo* hepatic phenotype
Subcellular fractions (S9, microsomes, mitochondria)	• Easily availabe depending on cell source used • Easy to use • Mitochondrial studies • Rapid metabolomic studies	• Absent cell structures/processes • Technical isolation procedures • Poor extrapolation to *in vivo*
Hepatocyte like cells (iPSC-derived hepatocytes	• Capacity to model patient diseases • High self-renewal capacity • High metabolic competence relative to immortalised cells	• Differentiation protocol influences terminal competence (variability) • Immature hepatic phenotype

HepG2 cells, a continuous cell line derived from a hepatocellular carcinoma, are considered viable for long term studies due to their ease of application while retaining some of the metabolic properties of PHH. HepG2 cells are consistent in their gene expression profile in early passage which is valuable as it allows for experimental reproducibility and robustness [81]. However, concerning aspects of the HepG2 phenotype include the notably lower expression profile of certain phase I enzymes such as CYP2D6 with a relative abundance up to 4.3% of the hepatic pool and ~20% contribution to commercial drug metabolism [69, 81]. Contrasting the poor CYP2D6 abundance is the overexpression of other enzymes such as CYP2W1 with the 150-fold increase in expression compared to pooled PHH. Some enzymes, such as CYP19A1, are found in HepG2 cells but are not expressed in pooled PHH [82]. These discrepancies highlight the alterations of HepG2 cells, which could potentially be manipulated *in vitro* to better phenotypically resemble PHH. Published data suggests an improved expression of several cytochrome subfamilies, promoting a more competent hepatic phenotype, when HepG2 spheroids are cultured past 21 days after seeding [83].

Historically, most of the hepatotoxicity screening has used 2D cell culture models. This raises several concerns, specifically when comparing the toxicity of a compound tested using 3D counterparts *in vivo*. It was previously believed that monolayer cultures could be used to reflect the approximate physiology of *in vivo* tissue [84]. However, when considering the physiologically incompatible nature of hard flat culture surfaces the opposite is true. Synthetic surfaces with reduced extracellular matrix components and limited cell-cell adhesion do not mimic the natural cell environment. Adhesion in 2D cultures is limited to that existing between neighbouring cells which is distorted by adherence to the culture vessel. Consequently, cell cultures lose the ability to mimic the classical *in vivo* cell-cell communication pathways established through gap junctions and chemical paracrine cell messaging systems [84].

The implications of this artificial cellular microenvironment are that cultures lose tissue related functions which hinders their predictive power. As an example, PHH in 2D cultures can dedifferentiate and die within 3-4 days. Continuous cell lines are considered more robust in maintaining their phenotype. However, these cells are also prone to varying states of differentiation and have reduced metabolism in comparison to primary tissues. This is an especially important consideration in metabolomic based toxicity screening [78].

1.8. 3D culture formats

When considering research under the umbrella of precision medicine, the success rates in drug development are expected to improve proportionally to the emergence of new technologies. The discovery of novel biomarkers has given rise to more precise drug targets and new preclinical models that are reported to better recapitulate *in vivo* biology and microenvironmental factors. One such technique which has helped bridge the gap between *in vitro* and *in vivo* systems is the adaptation of traditional monolayer cell culture techniques into 3D cultures. It is now well-accepted that culturing cells in 3D systems that mimic key factors of tissue is much more representative of the *in viv*o environment than simple 2D monolayers [84, 85]. While traditional monolayer cultures are still predominant in cellular assays used for high-throughput screening (HTS), 3D cell cultures techniques for applications in drug discovery are making rapid progress [86-89].

Currently, a wide variety of both static and fluid 3D culture techniques are available for culture of human hepatocytes. Static techniques are divided into hydrogels and scaffolds which can be synthetic or natural in origin as well as scaffold free cultures [90].

The advantages and limitations of several methods are described in Table 1.4 [90].

Table 1.4: 3D cell culture method advantages and limitations

Culture technique	Description	Advantages	Limitations
Static suspension cultures	Cells seeded in a low attachment material and maintained in suspension in media, spheroids form from cell aggregation	• Relatively low cost • Cell type dependent success • High throughput • Ease of application	• Continuous passaging is difficult • Inconsistent spheroid formation
Bio-scaffold cultures	Cells grown on a natural, isolated bio-scaffold such as collagen or basement membrane	• Promotes cell-cell adhesion • Allows for extracellular matrix adhesion • Cells regain morphology • If executed correctly is more representative of culture environment	• Difficulties in sample collection • Growth factor interplay with isolated basement membrane, poorly standardized, expensive, inconsistent spheroid size/ generation
Microfluidic device	Plastic chip array with inlets for media and specialised wells/inlets for spheroid formation	• Highly controlled spheroid size, ensures uniformity, perfusion of media supporting nutrient / waste exchange	• Culture harvesting can be difficult • Constricting culture environment • Specialised equipment required
Spinner/bioreactor dynamic suspension cultures	Cells seeded in culture vessels and grown under constant rotation, spheroids form via aggregation	• Mass production • Long term maintenance • Less media exchange required • Simple execution	• Require costly specialised equipment • Spheroid uniformity difficult to control • Inconsistent spheroid growth
Synthetic scaffold systems (agarose)	Cells grown using agarose as scaffold for spheroid formation	• Promotes cell-cell adhesion, cost effective • Promotes compact spheroids	• High density scaffold limiting cell proliferation • Synthetic surface • Tedious sample collection
Magnetic levitation	Cells treated with magnetic nanoparticles and held in suspension using neodymium magnets	• Relatively simple implementation once established	• Requires special equipment/ media • Colorimetric interference of media within magnetic fields causes inconsistent spheroid formation
Hanging drop	Cells grown in medium alone with aggregation occurring at apex of drop	• Control over spheroid size • Relatively low cost, generation of uniform spheroids • Adaptable to high throughput	• Drop variations influence cell kinetics, drops prone to fall • Tedious medium exchange • Expensive if using specialised plates

Static cultures are good for long term maintenance of liver specific functionality, have an increased sensitivity towards drugs and support long term expression of phase I and II enzymes. However, lack of standardization and difficulty in recovery of cells in tethered cultures (biogels or scaffold) remains problematic [91]. Fluid or dynamic culture systems are advantageous in that they allow for drug flow, rapid differentiation of cells, sustained functionality and increased liver-specific functionality. Shortcomings include low throughput and a lack of standardization [91].

1.9. Role of the extracellular matrix in cellular processes

Multicellular organisms require a structural framework to provide tissue integrity and define tissue boundaries. Initial studies on the intricacies of the extracellular matrix (ECM) and how its organization influences cell behaviour were conducted by Mina Bissell and her team in the 1980's [92]. It is now well known that ECM proteins are essential for the preservation of cell viability and to promote tissue specific phenotypes, this is especially true in the case of hepatocytes. The differential impact of selected matrix proteins (collagens, fibronectin, and laminin) has been previously described in monolayer cultures [93, 94], and more recently in 3D spheroids [95].

The ECM is a fibrous, hydrogel containing network consisting of proteins, proteoglycans, and glycosaminoglycans arranged into tissue specific 3D conformations to provide cells with topographical structure and signalling cues. It is a dynamic matrix to which cells attach and that enables cell migration [96] through growth factor and morphogen sequestration and release [97, 98]. The ECM provides binding sites for various cell surface receptors such as integrins. These integrin-ECM interactions can trigger cell signalling cascades regulating survival, proliferation, differentiation, and even apoptotic cues [99-101]. The function of the ECM is dynamically adaptable, controlling tissue texture, playing a central role in embryonic development and wound healing, and has a differential contribution to tissue-specific microenvironments [102]. Crosstalk within these microenvironments exist, with

resident cells producing and altering their surrounding ECM while simultaneously receiving signalling cues defining their own function.

The evolving understanding of the ECM has contextualised its importance in bridging pitfalls in the relevance of preclinical screening models. Figure 1.6 represents the comparative role of the ECM within 2D and 3D cell cultures and highlights its involvement in signal transduction.

Given the essential contribution of the ECM to maintain phenotypically relevant cell states, and the dynamic repurposing-remodelling relationship with the cells, assessing the phenotypic changes of cells when transitioning from an environment relatively devoid of ECM (2D monolayer cultures) to one which enables cellular deposition of ECM (scaffold free 3D spheroids) would be worthwhile to investigate.

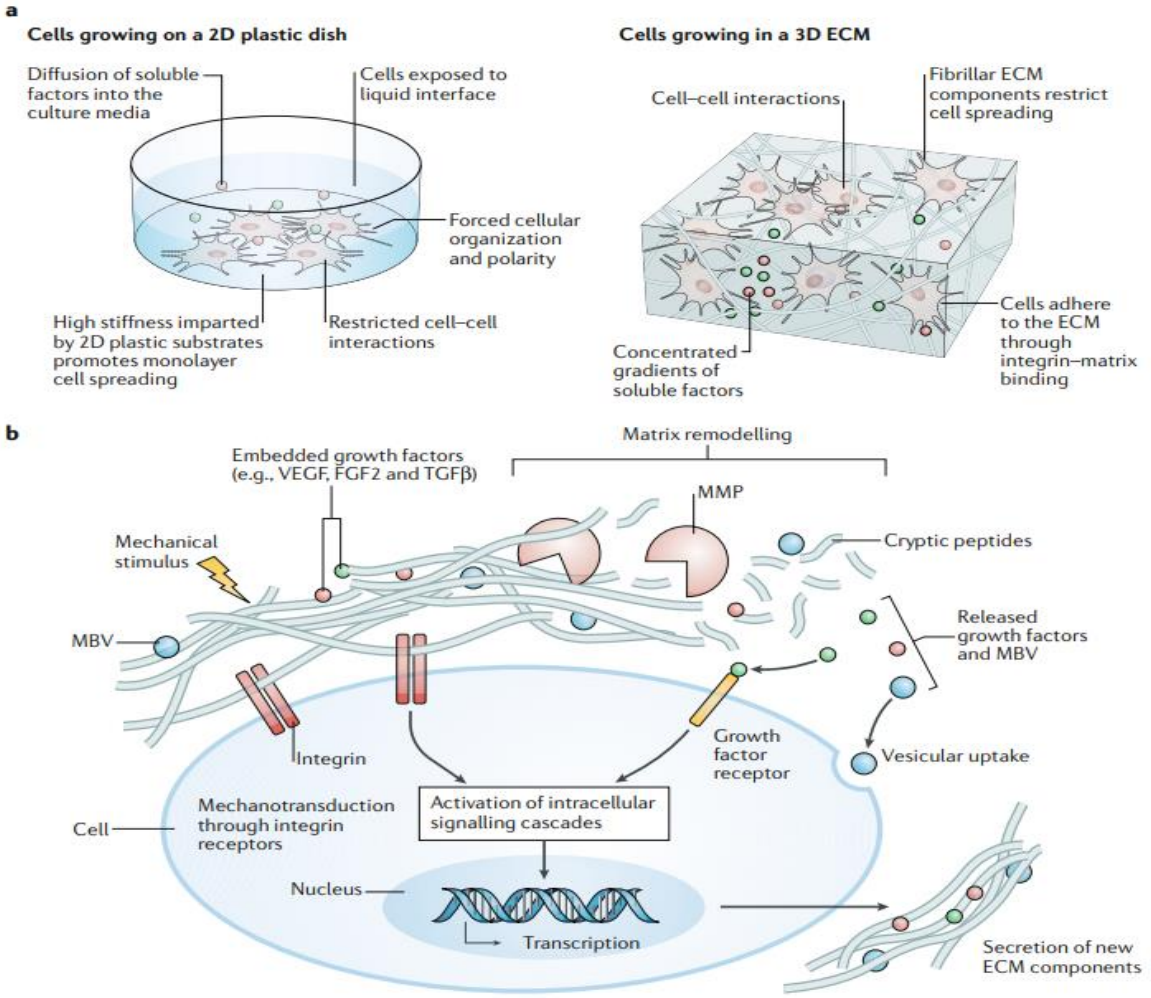


Figure 1.6: Cell–extracellular matrix interactions and matrix remodelling. A. 2D plastic substrates (top left) restrict cell attachment in a planar direction and force cells into an apical–basal polarity but allow diffusion of secreted soluble factors in the culture medium. By contrast, 3D extracellular matrix (ECM) substrates (top right) enable cell attachment in both planar and perpendicular directions without restricting cell polarity. The discrete matrix fibrils sterically hinder the spreading of cells and contribute to the sequestration of secreted growth factors into concentration gradients within the matrix. B. Cells interact with their environment through integrin and growth factor receptors. Cells convert mechanical stimuli from the ECM into biochemical activity through the binding and activation of integrin receptors, resulting in the activation of intracellular signalling pathways, activation of gene transcription and synthesis and secretion of ECM components. During matrix remodelling, proteolytic degradation induced by matrix metalloproteinases (MMPs) results in the release of tethered growth factors and matrix-bound nanovesicles (MBVs), as well as the production of cryptic peptides. The released bioactive components can interact with cells to promote diverse cellular functions such as proliferation, migration and differentiation. FGF2, fibroblast growth factor 2; TGFβ, transforming growth factor-β; VEGF, vascular endothelial growth factor. Taken directly from [103] with permission from Springer Nature, licence number 5452981407455.

1.10. The omics

The emergence of the omics technologies has revolutionised the understanding of systems biology. Genomics, the first of the omics technologies, focused on interrogating single genes or gene variants and provided a useful framework for the assessment of the underlying mechanisms contributing to hereditary and complex genetic disorders [104]. Since then, a there has been a rapid expansion in the emergence of new omics technologies which have all made significant intra-omic improvements proportional to the concurrent increases in computing power. Omics technologies include genomics, epigenomics, transcriptomics, proteomics, metabolomics, lipidomics and microbiomics. In the past decade, the use of combined omics approaches has provided a multi-faceted understanding of the maintenance of cell systems, developmental biology, and disease progression. An important consideration is that the foundation of omics technologies follows an extension of the central dogma of biology, in that, simply, the understanding of genomics is essential for the understanding of transcriptomics and the understanding of transcriptomics is essential for the understanding of proteomics, and so on as represented by Figure 1.7. The correct correlation and application of these is essential in multi-omic studies.

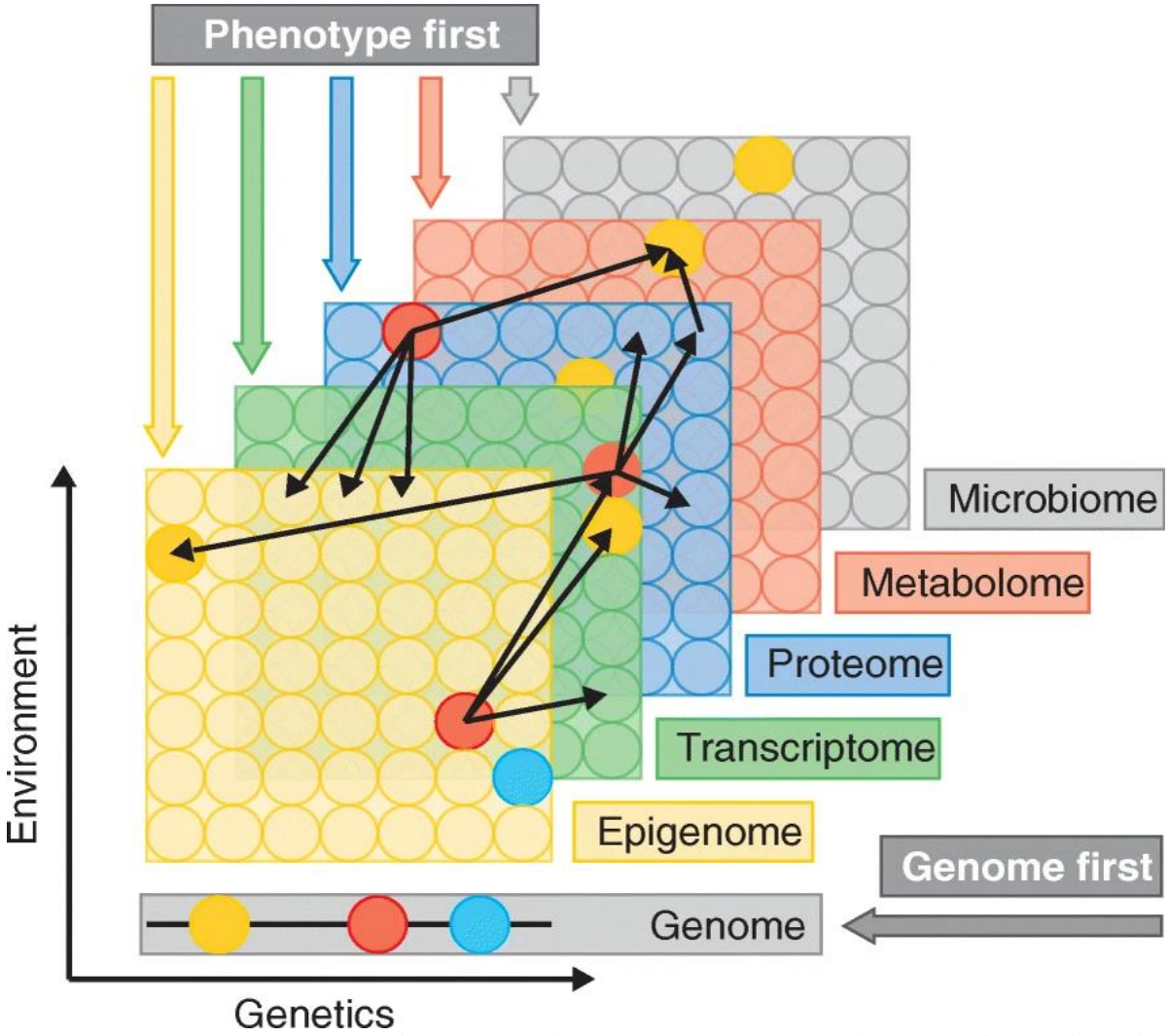


Figure 1.7: Interrelatedness of multi-omics data. Circles represent the collected data for that form of omics technology resulting from an underlying contributor on the genome. Axes represent the contribution of genetic underpinnings and environmental factors. Arrows represent the interrelatedness of data from one omics platform to another. Image used from [104] with permission from BMC under the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

1.10.1. Proteomics

The term "proteomics" refers to the systematic study of all proteins present in a cell or tissue at a given time. This includes describing their structure, function and expression in various biological systems [105]. Proteins are essential to all biological activities and their localisation, identification and quantitation provides insight into complex cellular regulatory networks [106]. Before the genomics revolution, chemical methods such as the stepwise Edman N-terminal degradation were used to determine

the sequences of single, highly purified protein samples. This was time consuming, costly, large enough samples were difficult to obtain and in certain instances sequencing was impossible to replicate [107]. With the emergence of mass spectrometric protein analysis and protein sequence prediction from genome sequences, large scale protein identification by correlating mass-spectrometric data with the available translated gene sequence databases [108] became routine. Proteomic studies can now reliably and reproducibly determine diverse protein properties including the amino acid sequence, abundance, post-translational modifications, protein-protein interactions, location and structure [109].

Whole proteome analysis remains challenging because of its vastness and unknown complexity. The number of genes in a species is not representative of the number of proteins present in an organism's phenotype. Protein numbers are higher in comparison, due to alternative splicing, RNA editing and degradation, post-translational modifications as well as protein complexing. Moreover, the range of protein concentrations generally exceeds the dynamic range of any single analytical method. High-throughput proteomic analysis workflows generally consist of four stages: sample extraction and preparation, protein or peptide separation, mass spectrometric (MS/MS) analysis, and protein identification inference and quantitation using bioinformatics tools. Common approaches in proteomics include intact protein sequencing using MS (top-down proteomics), separation at the protein level by one or two dimensional gel electrophoresis which can be followed by proteolysis and mass spectrometry (GeLCMS), and MS analysis of peptides following enzymatic or chemical cleavage of proteins (bottom up or shotgun proteomics) [110]. The shotgun strategy allows high-throughput identification of thousands of proteins from highly complex mixtures such as cell lysates [111].

Proteomics analysis can be targeted or non-targeted with non-targeted approaches relying on matching of peptide fragmentation spectra from individual precursor peptides collected using specific mass spectrometer acquisition parameters, which often confers a bias for more abundant peptides. Targeted proteomics differs

fundamentally in that the mass spectrometer is programmed to analyse only a preselected group of unique peptides using selected or multiple reaction monitoring (SRM/MRM) and is used to quantitate specific proteins.

1.10.2. Tandem mass tagging

In the field of proteomics, there are a vast array of techniques available for the isolation and identification of proteins. The mainstay of protein profiling was the use of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) to separate highly complex protein mixtures followed by spot isolation, reduction, alkylation, proteolysis to peptides and finally mass spectrometric peptide sequencing [112]. This process is laborious, shows poor reproducibility and is limited in dynamic range [113]. This results in poor correlation between sample replicates and causes difficulties in performing quantitative analysis between samples.

For these reasons, selection of the appropriate technique for proteomic analysis and peptide identification is crucial for generating reproducible and reliable data. Tandem mass tag (TMT) based isobaric tagging (or isobaric tags for relative and absolute quantitation: iTRAQ) enables samples under different conditions to be multiplexed and combined for relative quantitation by mass spectrometry. Reagent sets are available in various multiplexing options capable of labelling 2, 4, 6, 8 or 10 comparable samples obtained from cells or tissues which can be analysed simultaneously. Each mass tagging reagent has the same nominal mass (equal Dalton weight) chemical structure and chemical structure which is composed of an amine-reactive NHS-ester group, a spacer arm or mass normalizer, and an MS/MS reporter (Figure 1.8). For each sample, the unique reporter in the low mass region of the MS/MS spectrum (126-131Da for TMT6 Isobaric Label Reagents) is used to infer relative protein expression during peptide fragmentation [114]. Isobaric labelling exhibits a wide dynamic range in profiling both high- and low-abundance proteins. It can be used to identify and quantify proteins across a wide array of physiological properties

with diverse molecular weight and pI ranges, functional categories, and cellular locations [115, 116]. TMT can be employed in targeted MS analysis to track specific peptides or proteins of interest.

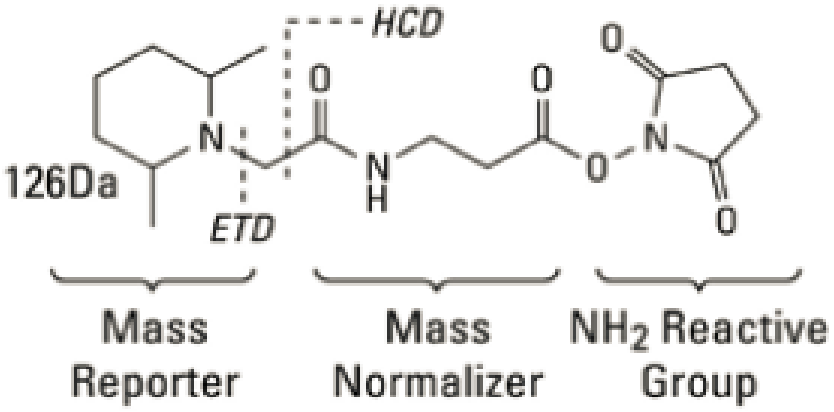


Figure 1.8: Representation of the general chemical structure of an isobaric TMT tag having different regions with stable isotope atomic distribution to provide a series of fixed mass molecules that provide different reporter mass fragments during fragmentation in the mass spectrometer (126 Da in this example).

Inherent limitations of both iTRAQ- and TMT-based quantifications include the requirement to observe low m/z fragment ions (reporter ions) and accurately resolve the mass difference (less than or equal to 1 Da), which limits the type of mass spectrometers that can be used. Because iTRAQ and TMT label-based quantifications are measured at the MS/MS level, potentially higher signal/noise ratios may be obtained for quantification compared with those obtained at the MS level. Additionally, the accuracy of MS/MS level quantification depends on the isolation window for selected precursors in the first stage MS, as all ions within that window will fragment, and potential interferences could skew the quantification results [117].

1.10.3. Bioinformatics for proteomics

The successful development of high-resolution MS-proteomics has hinged on concurrent developments in computational science and computing power. Together these systems have been at the forefront driving advances in systems biology. Protein

identification requires the matching of peptide level MS/MS spectral data with reference libraries either generated in previous experiments or by theoretical fragment ion predictions by specific computational search engines aided by neural networks [118, 119]. Following library scanning, peptide candidates are assigned confidence scores based on mathematical algorithms weighing a variety of search engines, user-defined and workflow specific characteristics. Confident peptide spectrum matching (PSM) is essential for downstream protein assignment, and the selection criteria for peptide-protein assignment is central to experimental validity. Algorithmic variances across search engines in peptide assignment can be a limitation and may introduce engine-specific biases in identified protein cohorts. The use of multiple search engines simultaneously can circumvent this issue but carries the potential of increasing the false discovery rate (FDR) [120]. Once initial identification and quantification workflows are completed, a new challenge is to provide functional interpretation of proteomic datasets. Successful analysis of proteomic datasets seeks to properly annotate functional or clinical implications, a process that hinges on proper sample data representation, which requires adequate and applicable data clean-up, normalisation, and statistical analysis. In the case of large proteomic datasets, statistical testing for significant differential expression is confounded by multiple hypothesis-testing that results in an increase in type I error frequency (incorrectly identifying significant differences) making the correct choice of statistical correction method essential to delineate these errors.

Functional assignment of correctly processed proteomic datasets is largely driven by gene ontology (GO) hierarchy as developed by the GO consortium. The GO vocabulary is organised into a directional, acyclic, tree like node structure where higher order terms are broken down into multiple constitutive lower order nodes with increasing annotation specificity [121]. Gene ontology is also useful to provide probable function of protein subsets when assigned according to gene inferences for similar sets of proteins, for example, annotation of function, subcellular localisation and biological pathway involvement is useful in assigning protein function with a

degree of statistical confidence. Figure 1.9 represents a typical pipeline for the analysis of proteomic data.

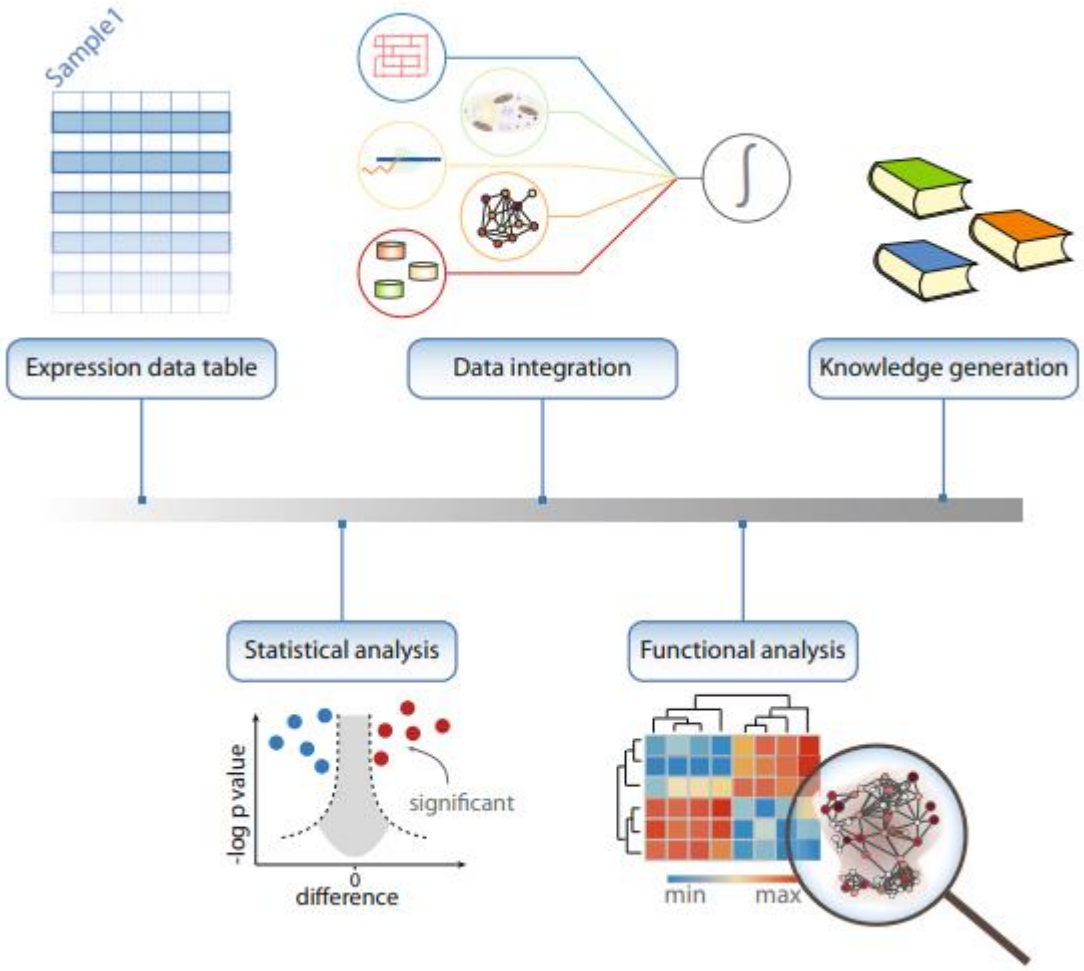


Figure 1.9: Example of typical proteomic data analysis workflow, taken from [122] with permission from Springer Nature under the terms of the Creative Commons CC BY license.

1.10.4. Metabolomics

Metabolomics is another omics field related with the study of metabolites, which exploits a combination of sophisticated analytical chemistry techniques and advanced computational methods to characterise complex mixtures. Metabolites represent one of the downstream outputs of the genome and a product of the environment. Therefore, metabolomics may be used to explore the nexus of gene-environment interactions and by extrapolation, provide protein-metabolite interactions. A significant application of metabolomics is in the analysis of how drugs are

metabolised, providing data for the regulatory requirement of identification of the metabolic products of new drug candidates. Additional information is in ascertaining which enzymes are involved in specific metabolic pathways as well as flagging potential concerns of drug-drug interactions and pharmacogenetic idiosyncrasies.

Widely used techniques in metabolomic research include nuclear magnetic resonance (NMR) spectrometers and gas/liquid chromatography-mass spectrometry (GC/LC-MS). Other techniques such as electrochemistry and infrared spectroscopy have not been widely adopted due to being bulk property analysis techniques with poor specificity, low resolution and lack of structural elucidation [123]. Given the disparity in analytical coverage of widely used metabolomic techniques these applications are used either individually or in combination for targeted or shotgun metabolomics, each leveraging specific advantages. Generally, NMR provides good metabolite identification, quantitation and automation but sacrifices sensitivity and carries a greater start-up cost compared to MS applications. GC-MS is robust with good sensitivity and convenient identification of common volatile metabolites from commercial databases, but sample preparations can be labour intensive, and identification of novel compounds is challenging. LC-MS/MS is the most widely adopted method and confers high sensitivity, wider mass detection ranges and coverage of soluble compounds and easily amenable to different methodologies, however it is less robust [124-126].

1.11. Study rationale and aim

As the implementation of spheroid cultures becomes more mainstream in hepatic preclinical *in vitro* modelling, there is a requirement for better understanding of the phenotypic alterations conferred to these model systems. HepG2 cells remain central as an initial hepatotoxicity screening platform and to this end have also been adapted for use in spheroid culture models. However, given the multitude of spheroid generation methods as well as the variety of functional endpoints required by

preclinical screening models, the degree of their usefulness has not been established robustly. There is also little understanding regarding the temporal proteomic changes occurring within these systems and the time points at which optimum recapitulation of cellular phenotypes is achieved. This research set out to contextualise any phenotypic changes as modulated by culture conditions as well as the effect of external environmental stimuli relevant to the liver i.e., extended exposure to a low concentration drug cocktail.

1.11.1. Aim

The aim of this study was to compare the observable proteomic changes of HepG2 cells cultured as 3D spheroids over time and after induction with a prototypical drug cocktail using stable isotope labelled tags for relative quantitation and mass spectrometry-based proteomic methodology.

1.11.2. Objectives

1. To use the sulforhodamine B assay on HepG2 monolayer cultures to determine the maximum sub-toxic concentration of individual drug compounds from a drug cocktail containing Phenacetin, Diclofenac, Omeprazole, Dextromethorphan, Midazolam, Buspirone and Artemisinin
2. To optimise and characterise the cell culture parameters for long term HepG2 spheroid culture
3. To culture viable HepG2 cell spheroids in the presence of the predetermined sub-toxic concentrations of the induction drug cocktail
4. To perform microscopic evaluation of spheroid cultures over an extended time course (0 – 28 days) using phase-contrast microscopy and live-dead staining
5. To confirm the lack of cytotoxicity of the induction drug cocktail in spheroids using the lactate dehydrogenase cell enumeration assay
6. To develop and optimise a confocal microscope method for visualisation of selected hepatic function marker-protein expression and spheroid ultrastructure
7. To generate protein mass profiles of protein samples using SDS-PAGE to confirm protein extraction
8. To characterise the temporal proteomic changes occurring in HepG2 spheroids

9. To determine the effects of extended drug exposure of specific metabolic enzymes on HepG2 spheroid cultures over time by means of stable isotope labelled tags for relative quantitation by mass spectrometry-based proteomics
10. To assess metabolic competence of spheroid cultures using LC-MS/MS method following a probe drug challenge

Chapter 2. HepG2 spheroid culture characterisation

2.1. Brief overview

Hepatic cell culture models are essential tools in the preclinical phases of drug development. Assessment of cell growth kinetics, the expression of relevant marker proteins and dose related cytotoxicity profiles are important prior to undertaking long-term studies using these models. The wide availability and ease of culturing of HepG2 cells has resulted in their routine use in hepatic cytotoxicity screening of drug candidates and existing drugs. However, as with all immortalised cell lines, functional attenuation, dysregulation of phenotypic characteristics, and genetic aberrations are all phenomena which occur proportionally to time in culture and exposure to repeat stress events [127]. Additionally, while these cells have been well described in literature with regard to their use as in monolayer culture [128-131] the reports of their adoption as 3D type spheroid cultures is still sparse. The numerous spheroid generation and maintenance methods all contribute unique caveats when comparing methods. For example, the use of matrix or scaffold supplementation to aid in cell adhesion may induce variable drug responses due to drug-matrix interactions. Biological matrices also have the potential to mediate cell signalling pathways which may be otherwise absent in scaffold-free systems [132]. Additionally, long term studies on spheroid models require the maintenance of viable cultures maintenance of similar size spheroids over the specified culture period and all the parameters enabling these need to be considered.

2.1.1. Selection of cocktail drugs

Induction of hepatic enzymes is an adaptive response associated with liver enlargement, induction of gene expression, and changes in cellular conformations [133, 134]. Induction is triggered by ligand activation by xenobiotic ligand compounds, with several nuclear receptors (PXR, CAR, and AhR) that upon ligand

binding, translocate to the nucleus where together with various co-activators, bind to response elements that then mediate trans-activation of target genes and expression of drug metabolising enzymes. An overview of the process is provided in the context of phenobarbital mediated induction of CAR (Figure 2.1).

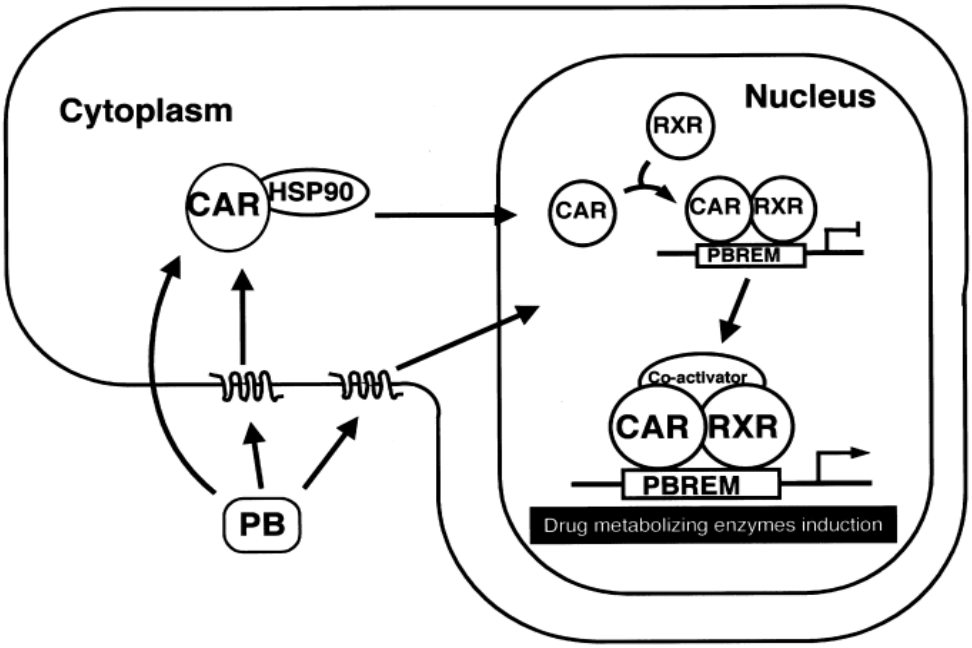


Figure 2.1: Schematic representation of CAR activation by phenobarbital (PB). Cytosolic CAR sequestered by heat shock protein 90 (HSP90) is activated by PB binding and translocates to the nucleus where it heterodimerizes with retinoid X receptor (RXR). The heterodimer then binds to enhancers of target genes, in this case PBREM, which is an enhancer for CYP2B genes. The complex together with other co-activators induce the expression of drug metabolizing enzymes [135]. Reprinted with permission from Taylor and Francis, reference number: iemt/02836149.

An extensive number of studies exploring the mechanisms of CYP enzyme induction have been reported [135-139], however, the process is still poorly understood. Multiple hypotheses exist regarding alternate mechanisms for CYP induction which may underlie or act concurrently with those described, such as, the interplay of unknown receptors and/or orphan receptors with undescribed cofactors [135]. The above-described interaction of nuclear receptors with various xenobiotics implies that simultaneous induction of multiple families of CYP enzymes is possible, and this has been described previously [140, 141]. Broad-spectrum or prototypical inducers such

as carbamazepine have been shown to concurrently upregulate CYP 1A2, 2B6, 2C9, 2C19 and 3A4/5 families. Coupled to the induction or inhibition of their metabolizing enzymes, various xenobiotics have been shown to up- or down-regulate other proteins, as in the case of the upregulation of multidrug resistance-associated protein 2 (MRP2). Careful consideration is required when undertaking drug exposure studies as not all xenobiotics induce or suppress enzymes during exposure, and some may do so even when predicted not to through interaction with expression regulators. These phenomena are likely due to the intrinsic properties of each drug and the way the individual chemical structures interact with different receptors, and the biological context of the hepatocytes mediating this metabolism. A drug cocktail (Table 2.1), validated as substrates for hepatic enzymes, was selected based on the criteria that the compounds are clinically approved and have limited potential for hepatocytotoxicity when used as recommended physicians in patients without genotypic or pathophysiological liver impairment.

Table 2.1:Details of drugs selected for potential toxicity drug cocktail and CYP selective metabolism assays

Drug Name	Drug Class	Nuclear receptor binding	CYP Class * Inducer # Inhibitor ^ Substrate	References
Artemisinin	Antimalarial	PXR, CAR	CYP2B6*^ CYP3A4*^ CYP2C19*	[142, 143]
Buspirone	Anxiolytic	N/A	CYP3A4^ CYP3A5^ CYP3A7^ CYP2D6^	[144, 145]
Dextromethorphan	Antitussives	N/A	CYP2D6^ CYP3A4^ CYP3A7^ CYP2B6^ CYP2C19^ CYP2C9^ CYP2B4^	[146-151]
Diclofenac	NSAID	N/A	CYP2C9^ CYP2C19^ CYP1A2^ CYP2C8^ CYP3A4^# CYP2B6^ CYP2C18^	[152-156]
Midazolam	Benzodiazepines	N/A	CYP3A4^# CYP3A5^ CYP3A7^	[157-162]
Omeprazole	PPI	AhR	CYP1A1*^ CYP1A2* CYP1B1* CYP2C8^ CYP2C9^ CYP2C18^ CYP2D6^ CYP3A4^*# CYP2C19^#	[163-169]
Phenacetin	NSAID	N/A	CYP1A2^ CYP2D6^ CYP1A1^ CYP1A13^ CYP2A6^ CYP2C19^ CYP2C9^ CYP2E1^ CYP3A4^	[170-174]

2.2. Materials and Methods

All reagents and consumables used in this study were purchased from reputable vendors and used within the specified shelf lives. Reagents and consumables used in cell cultures were either purchased as sterile or sterilised using an autoclave (121°C, 45 min). Consumables for cell cultures were all purchased from Gibco (Thermo Fisher, Waltham, MA, USA) unless otherwise stated. All percentage dilutions were as weight per volume (w/v) unless otherwise specified. All experimental assays were conducted as 3 technical replicates with 3 biological replicates unless otherwise specified. Cell cultures were thawed at passage 12 and maintained no further than passage 35 with biological replicates of cell cultures no more than 5 passages apart. Fresh stocks of drug cocktails were prepared and used throughout biological replicate sets. Routine mycoplasma PCR screening was conducted every month, and cultures were visually monitored daily for other sources of contamination or abnormalities. All proteomic centric workflows made use of Protein LoBind® Tubes (Eppendorf; Hamburg, Germany) to avoid selective loss of protein groups and maximize sample recovery.

2.2.1. Monolayer cultures

Two-dimensional cell cultures were seeded and maintained under standard cell culture conditions (37°C with 5% CO_2, 90-95% humidity). HepG2 cells were obtained from Cellonex (Johannesburg, RSA: CHG2-C) and cultured in Dulbecco's modified minimum Eagle medium (DMEM) supplemented with 10% foetal bovine serum (FBS), and 2 mM GlutaMAX. Cells were cultured in various culture vessels at the appropriate densities. Densities when plating 2D cultures did not exceed 2×10^4, 5×10^5, 2×10^6, and 5×10^6 cells per 96-well plate, 6-well plate, 25 cm^2 or 75 cm^2 flask respectively. Cells were maintained until 80-90% confluence prior to passaging or harvesting using recombinant dissociation enzyme TrypLE for 5 min at 37°C. Culture media was added, and dissociated cells were collected and transferred to a sterile Falcon tube, centrifuged (200 *g*, 5 min) and counted using the trypan blue (0.1%) exclusion assay on a Countess II Automated Cell Counter (Invitrogen, Waltham, MA, USA).

2.2.2. *Three-dimensional culture generation*

Three-dimensional cultures, in general, have been adopted far less frequently than traditional monolayer cultures and carry an increased technical complexity for generation. Careful consideration is required with regards to application and study endpoints. Useful measures are the ease of applicability and maintenance, adaptability to high throughput screening (HTS), biological relevance and potential interference of culture scaffolds with assay parameters. In some applications, introduction of an external biological scaffold may allow for a more representative cellular architecture to be achieved [175], however batch to batch variations, inconsistencies in formation kinetics and interference with assay readouts, present variables that are difficult to control for. For these reasons, two spheroid generation methods were selected for this study. The hanging drop method was selected as the most appropriate method for downstream whole-cell proteomic assessments as it provided a static, scaffold-free system of spheroid generation. Perfecta3D® 96-well plates were purchased from 3D Biomatrix (Michigan, USA) and culture conditions optimised for seeding density, growth kinetics and medium volume in a previous study [176]. For metabolomic and microsomal assessments, the 3D Petri Dish® method was deemed most appropriate due to its ease of application, ability to generate a large number of spheroids per plate and superior scalability for HTS. The 3D Petri Dish® method was selected and 9x9 spheroid micro-moulds, suitable for the wells of a 12-well plate were purchased from Microtissues®.

2.2.3. *Generation of hanging drop spheroids*

Hanging drop plates were prepared by addition of autoclave-sterilized molten agarose (1%) to the upper and lower plate reservoirs to aid in humidity control. HepG2 cells were harvested and counted as described above, seeded at 20 000 cells/well in 45 μL of culture media and incubated at 37°C under 5% CO2, 90-95% humidity. Spheroid formation occurred by cellular aggregation due to gravity at the

apex of each drop over several hours. Spheroid cultures were maintained for up to 42 days with partial media exchange (10-12 μL) every alternate day. Spheroids were harvested by aspiration using an air displacement pipette.

2.2.4. Generation of 3D Petri Dish spheroids

Micro-mould 3D Petri Dishes were rinsed with deionised ultrapure water (dH_2O) before autoclave sterilisation. Agarose for casting was prepared by dissolving 2 g of agarose powder in 90 mL of phosphate-buffered saline (PBS), pH 7.4, and then topping up to 100 mL with PBS before autoclaving. PBS was used in favour of dH_2O to minimise any potential disturbance to in-well osmolarity due to leaching of dH_2O into the culture media. Moulds (9 x 9 array) were cast by the addition of 500 μl molten agarose (2%) into the mould reservoirs and allowed gel for 5-10 min. Once set, gentle flexing of the mould casing released the agarose mould reservoirs which was transferred to 12-well cell culture plates. In the cases where moulds fell in upside down, gentle reorientation was performed using a sterile pipette tip. Moulds were equilibrated by addition of 200 μL of culture media to the reservoir and plates placed in the incubator for 2 h until all the wells were free from bubbles. Equilibration media was aspirated off using a pipette and harvested HepG2 cells, counted as described above seeded at 81 000 cells/mould in 180 μL of culture media by slow addition of the cell suspension to the mould reservoir. Plates were incubated for 1 hour to allow cell suspensions to settle in the wells of the micro-mould before slowly adding 2 mL of culture media to each well around the micro-mould. Plates were incubated under standard conditions and media changed weekly by vacuum aspiration of depleted media and replacing as described above. When harvesting spheroids, moulds were lifted vertically, and spheroids displaced from wells by gently rinsing with culture media and collecting using a pipette.

2.2.5. *Overview of drug exposure for the assesment of various endpoints in spheroid cultures*

During this study HepG2 spheroids were exposed to different concentrations of drugs to meet specific assay endpoints. The concentration of individual drugs and the combined drug cocktails used for cell culture drug exposures can be separated according to the context of two different experimental endpoints.

The first endpoint is for the assessment of proteomic changes in response to long-term exposure to sub-toxic concentrations of a drug cocktail. The drug cocktail for these experiments is detailed in Table 2.1 above and was to optimise sub-toxic drug concentrations for long-term drug exposure to assess downstream proteomic changes by mass spectrometry (MS) based experiments. Results characterising the cytotoxicity profiles of this cocktail are reported later in this chapter and within the context of downstream MS based proteomic experiments reported later in Chapter 4. In general control spheroid cultures that were not exposed to drug cocktail are denoted by the suffix (C) and those which had undergone drug exposure receiving the suffix (I).

The second drug cocktail (the Geneva cocktail [177]) is a combination of probe drugs used for functional metabolomic assessment of selected CYP450 enzymes. This cocktail was developed and validated for use within clinical settings for patient phenotyping experiments. This drug cocktail and the defined metabolites will be discussed in Chapter 4 as part of the metabolomic competence assays.

2.2.6. *Spheroid protein quantitation*

Protein was quantified using the bicinchoninic acid (BCA) assay, which makes use of the reduction of Cu2+ to Cu+ and after complexing the BCA, produces a colorimetric change proportional to protein content, originally described by Wiechelman et al. [178]. Pierce™ BCA Protein Assay Kit was purchased from Thermo Scientific™ and assay carried out using an inhouse optimized, modified version of the manufacturer's instructions. Briefly, spheroids were collected as described above, pooled, centrifuged

(500 g, 10 min) and washed with 5 volumes of PBS. Spheroids were lysed using radioimmunoprecipitation buffer (RIPA: 10 mM Tris-HCL (pH 8), 1 mM ethylenediaminetetraacetic acid, 0.5 mM ethylene glycol tetraacetic acid, 1% Triton X-100, 0.1% sodium dodecyl sulphate, 0.1% sodium deoxycholate, 140 mM sodium chloride) containing cOmplete™ protease inhibitor cocktail (Roche)). Lysis was aided by ultrasonic disruption on ice (1500 W, 5 min) and mechanical disruption of pellets using a Dounce Homogenizer. Samples were cleared by centrifugation (16000 g, 10 min) and supernatants transferred to new Eppendorf tubes. Working solution of BCA reagent was prepared by combining reagent A (2% sodium carbonate, 0.16% sodium tartrate, 0.9% sodium bicarbonate, and 1% BCA; pH 11.25) with reagent B (4% copper II sulphate pentahydrate) at a ratio of 50:1. Assay calibration curves were generated for 8 calibration points (0 – 2 mg/mL) using bovine serum albumin (BSA) included in the kit. Dilutions for calibration points were made in RIPA buffer to serve as an internal control for potential colorimetric interferences. Samples and standards were combined with BCA working reagent at a 1:40 ratio to a total volume of 200 μL/well. Assay plates were covered in aluminium foil and incubated for 30 min at 60°C in the dark before quantification using a Bio-Rad iMark™ microplate reader at 560 nm.

2.2.7. Establishment of baseline cytotoxicity profiles

Paracelsus is regarded as the father of toxicology and pioneered several concepts that are still emphasized in modern practices. One of his more notable contributions is that all compounds are toxic, and that the dose distinguishes a poison from a remedy [43]. Following this narrative, it is important to establish toxicological thresholds for any compounds used in cell culture assays. Drug induction studies require long term repeat exposure to compounds of varying concentrations and deleterious or toxic effects to cell systems, as a result of the drug itself or its metabolites must be avoided. These are important parameters that need to be established before undertaking any

proteomic analysis that requires functional cell systems and the use of subtoxic concentrations is therefore essential.

2.2.8. Sulforhodamine B cell enumeration assay

Cell enumeration assays are useful tools for cytotoxicity screening as increases or decreases in cellular densities may be used as an inference for growth or drug induced cytotoxicity. Changes in cell densities can be compared to negative controls to assess potential stimulatory, inhibitory, or toxic effects. Cell enumeration using a modified version of the Sulforhodamine B assay (SRB) described by Vichai *et al.* [179]. HepG2 cell monolayers were seeded at 20 000 cells per well in 100 μL of culture media. Cells were allowed to attach overnight and then exposed to eight half-log concentrations of each drug from 200 μM. Drugs were prepared as 50 mM stocks in dimethyl sulfoxide (DMSO) and diluted as required immediately before exposures. A vehicle control of 0.5% DMSO was included to correct for any effects contributed by the organic solvent. Cells were exposed to drugs for 72 h before fixation with 50% trichloroacetic acid (TCA) overnight at 4°C. After fixing, cells were washed gently under running water and placed in a dry heat oven at 60°C. Once dried SRB staining solution (0.057% SRB in 1% v/v acetic acid) was added to each well and cells stained for 30 min at room temperature. Plates were washed three times with 1% v/v acetic acid to remove excess unbound dye and dried as described previously. Bound dye was solubilised using 200 μL of pH 10 Tris-base solution (10 mM) under gentle agitation for 60 min. Absorbance was quantified at 510 nm with a reference wavelength of 610 nm using a Bio-tek Elx800 microplate reader.

2.2.9. Lactate dehydrogenase assay

Lactate dehydrogenase (LDH) is a ubiquitous enzyme present in a wide variety of cell types and released following damage to cell membranes. Free LDH present in cell culture media facilitates the conversion of lactate to pyruvate via the reduction of

nicotinamide adenine dinucleotide (NAD+) to NADH. The resulting NADH is then used as a cofactor for the reduction of tetrazolium salt to a formazan product (red) which can be quantitated using a plate reader at 490 nm. Formazan production is directly proportional to the amount of LDH present and can thus be used as an inference for cytotoxicity.

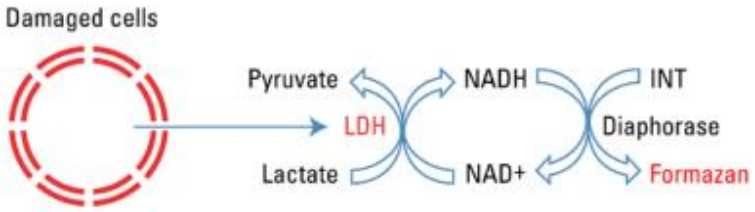


Figure 2.2: Schematic of lactate dehydrogenase assay mechanism.

Potential cytotoxicity in hanging drop spheroid cultures was assessed using the LDH assay. Spheroids were grown as described above as control or drug treated groups. Stocks of drugs from the cocktail described in Table 2.1 were prepared as described in Section 2.2.8 above and diluted in complete media to a final concentration of 5 μM. Cells for control spheroids were resuspended in complete media containing an equivalent solvent vehicle concentration (0.07% DMSO) only while drug treated spheroids were seeded into drug cocktail supplemented media. After select culture time points, LDH release from dead cells was measured using the Peirce LDH cytotoxicity assay kit (Thermo Fisher, Waltham, MA, USA) using a modified version of the manufacturer's protocol. Briefly, reaction mixture was prepared by combining 11.4 mL of assay substrate, solubilized in dH_2O, with 0.6 mL assay buffer under exclusion of light. Positive controls were prepared by diluting 1 μL of included LDH positive control with 10 mL of 1% BSA in PBS. Spheroids, (8 per hanging drop plate) were collected via air displacement, pulsed to sediment spheroids, and media transferred to a new tube and stored at -80°C for downstream quantification. Spheroids were washed 3 times with PBS and protein quantified as described above. For LDH maximum release controls (LDHmax), spheroids were collected as described

above and 40 µL of 10x lysis buffer was added to the tube containing media and spheroids (360 µL). Lysis was aided by mechanical disruption for 20 min before pelleting debris by centrifugation at 16000 *g* and collecting supernatant media. Sample collections, including controls, were done weekly from days 14 – 28 of the culture time courses. After completion of the time course, stored media samples were thawed and equilibrated to room temperature. Media from LDHmax controls was diluted 3x to ensure readings were within the linear range. LDH was quantified by combining 50 µl stored media from samples and controls with 50 µL reaction mixture to triplicate wells of a 96-well plate, incubated for 45 min and quenched by addition of 50 µL stop solution. Dual wavelength absorbance was measured using a Bio-Rad iMark™ plate reader at 490 nm and 680 nm. Percentage cytotoxicity was calculated by expressing absorbances of samples from each time point compared to the corresponding control and correcting for dilution of LDHmax controls. Potential differences in cell number were accounted for, and relative cytotoxicity values were normalised to quantified protein.

2.2.10. Light, fluorescence and confocal microscopic characterisation of spheroid cultures

2.2.10.1. Immunohistochemistry of monolayers and spheroids

Immunohistochemistry (IHC) tandem confocal microscopy is a useful technique for characterisation of protein expression and localisation within cell cultures. The principle of IHC utilises the highly selective binding of antibodies or selective ligands to specific antigens or receptors. This interaction is generally considered to be proportional to the antigen in the case of excess added monoclonal antibodies. The antibody (Ab) to antigen interaction can be visualised by introduction of various markers including conjugated fluorescent tags, conjugated enzyme-substrate colorimetric reactions, radioactive probe conjugation, and colloidal gold [180, 181]. However useful, without proper optimisation IHC may give rise to false positive, false negative, and non-specific binding results due to a variety of factors acting

concurrently or individually, such as, inappropriate fixation methods, inadequate antigen retrieval, use of incompatible mounting matrices, inefficient cell permeabilization, incubation times and temperatures, optical inaccessibility, and improper fluorescent stain concentration and/or primary/secondary Ab ratios. Additionally, to enable comparisons across experiments, coupled to the above-mentioned experimental parameters, several instrument parameters need to be optimised and remain constant during a set of assays. In the case of fluorescent confocal microscopy, variations in parameters such as laser strength, master gain, and aperture size across experiments make direct comparisons difficult. To enable tracking of protein expression changes over the time course, the parameters discussed above were optimised and kept constant across experiments. The details for antibodies used for IHC are present in Table 2.2. Spheroids were stained for various hepatic, structural and proliferation related proteins namely α-fetoprotein (AFP), albumin (Alb), cytokeratin 18 (CK18), and hepatic nuclear factor 4 alpha (HNF4α). These data will be presented here and in other chapters in the context of downstream proteomics data.

HepG2 cells were grown and harvested as described above and seeded at 60 000 cells/well onto autoclave sterilized (121°C, 45 min) 12 mm coverslips coated with 500 µl Corning® Matrigel® Growth Factor Reduced (diluted 1/30 in pre-chilled DMEM) in 48-well plates. Cells were maintained for 2-3 days before pre-fixing with 500 µl of 4% paraformaldehyde (PFA) for 10 min and then replacing with fresh 4% PFA with 8% sucrose for 40 min. Coverslips were washed 3 times with PBS before cell permeabilization and blocking in B-PBT (1% Triton X-100, 10% FBS, and 4% BSA in PBS) for 30 min. Samples were incubated with selected primary antibodies diluted in B-PBT (Table 2.2) for 2 h followed by 3 washes in B-PBT before incubating with secondary antibody (Table 2.2) for 2 h. Whole spheroids were harvested by pipette aspiration and fixed in 4% PFA with 8% sucrose overnight then washed 3 times for 1 h each in PBS on a plate shaker at 100 rpm. Permeabilizing and blocking was conducted as above in B-PBT for 2 h. Spheroids were incubated with selected primary antibodies in B-PBT overnight at 4°C followed by two washes for 2 h in 0.2% PBT

(0.2% Triton X-100 in PBS) and one wash in B-PBT for 2 h before incubating in secondary antibody (Table 2.2) in B-PBT overnight. Confocal images were acquired on a Zeiss LSM800 confocal laser scanning microscope as described in [182].

Table 2.2: Panel of primary and secondary antibodies for immunofluorescent staining

Primary antibody	**Species**	**Vendor**	**Dilution**
AFP	Mouse	Abcam	1/50
Alb	Chicken	Abcam	1/100
CK18	Mouse	Abcam	1/100
HNF4α	Rabbit	Abcam	1/100
Secondary antibody	**Species**	**Vendor**	**Dilution**
Alexa Fluor 488	Goat anti-chicken	Abcam	1/1000
Alexa Fluor 555	Donkey anti-rabbit	Abcam	1/1000
Alexa Fluor 647	Donkey anti-mouse	Abcam	1/1000

2.2.10.2. Spheroid size and morphology

Characterisation of spheroid growth and morphology are important metrics to define in 3D culture. In the context of hepatocyte spheroids, formation of well compacted, uniform spheroids has been shown to be an indicator of an optimal cell culture environment promoting more relevant ultrastructural organisation and phenotypic function [183]. Spheroid cultures maintained over a 28-day time course, were imaged on days 1, 3, 7, 14, 21 and 28 using a Zeiss AxioVert A1 with an AxioCam digital camera. Spheroids (8 per hanging drop plate, 10 per μmold) were collected and diameters, circumferences, and coefficients of roundness measured for each culture timepoint using the measure tool built into Zeiss Zen blue software package.

2.2.10.3. Live-Dead staining

Live-dead staining of spheroids allows for a two-colour discrimination of viable and compromised cell populations. Briefly, non-fluorescent membrane permeable fluorescein diacetate (FDA) is cleaved by the esterase enzymes present in viable cells, liberating the highly fluorescent fluorescein. Propidium iodide is a cell impermeable

dye which intercalates with the DNA double helix and can be used as an inference of compromised cell membranes of dead/dying cell populations. Spheroids were collected as described previously, centrifuged (200 *g*; 5 min) and washed with PBS. One millilitre of staining solution (10 µl of 2 mg/mL propidium iodide in PBS and 2 µL of 5 mg/ml FDA in acetone) was used to stain spheroids for 5 min at room temperature under exclusion of light. Spheroids were then washed 3 times with PBS and visualized using a Zeiss AxioVert A1 with an AxioCam digital camera using light emitting diode (LED) excitation with green or red fluorescence filter sets. Digital images were recorded and processed using the Zeiss Zen blue software package. Assay positive controls were spheroids exposed to 5 or 10 µg/mL puromycin for 16 h and prepared and processed in the same way were used as controls.

2.2.10.4. Expression, purification and validation of CNA35 collagen probes

The construction of a fluorescent collagen binding protein (CNA35) has been described previously by Krahn *et al.* [184]. CNA35 contains the soluble N1 and N2 collagen binding protein domains from S*taphylococcus aureus* bacteria [185]. Plasmids pET28a-mTurquoise2-CNA35 (Plasmid #61602) and [185] pET28a-tdTomato-CNA35

(Plasmid #61606) were purchased from Addgene (Watertown, MA, USA) and were transformed into XL10-Gold® ultracompetent cells (Stratagene).

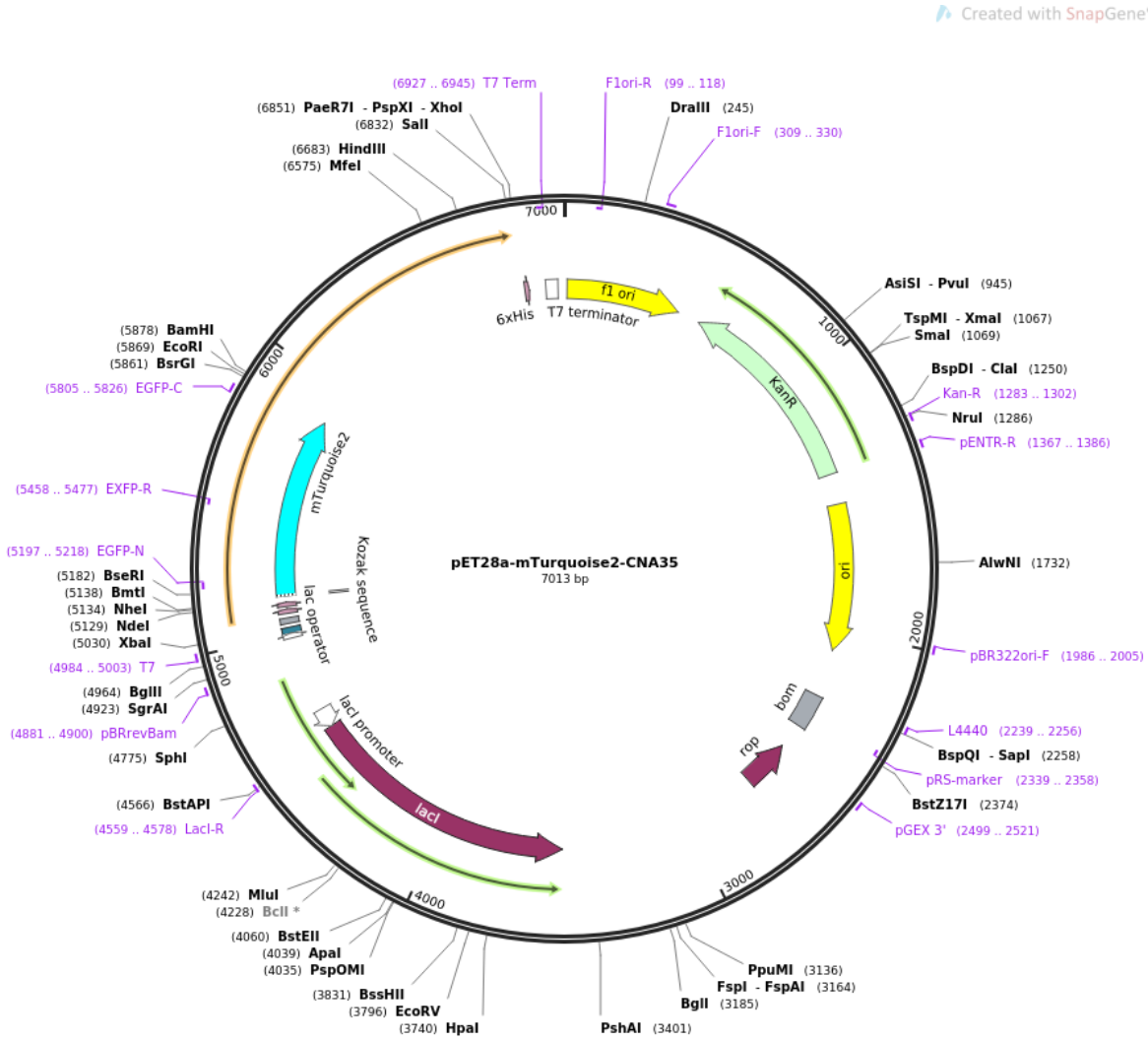


Figure 2.3: Schematic representation of pET28a-mTurquoise2-CNA35 plasmid used for expression of CNA35 containing Kanamycin resistance, lac operon, lac promoter and 6x histidine tag regions.

Single bacterial colonies were picked and used for inoculation of 8 mL Luria-Bertani (LB) broth cultures (10 g/L Tryptone, 10 g/L NaCl, 5 g/L yeast extract, pH 7.2) supplemented with 20 µg/mL kanamycin. Cultures were grown overnight in a shaking bacterial incubator at 37°C under circular agitation (250 rpm) until an optical density at 600 nm of approximately 0.6 was reached. Cultures were then used to inoculate a further 100 mL of fresh LB broth and maintained under the same conditions until a similar optical density was achieved. Lac operon mediated

expression of CNA35 was induced by introduction of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma-Aldrich). Protein expression was allowed to continue for 24 h before harvesting by centrifugation (6000 *g*, 10 min). Pellets were washed 3 times in PBS, resuspended and transferred to fresh Falcon tubes before lysing by ultrasonic disruption using a BioLogics model 3000 ultrasonic homogenizer (900 W, 30 second pulses). Bacterial lysates (1 mL) were transferred to 1.5 mL Eppendorf tubes and centrifuged at 16000 *g* for 15 min. The resultant soluble fractions were collected and used for purification. MagReSyn® NTA Screening Kits (ReSyn Biosciences, Gauteng, South Africa) were kindly donated by Dr Justin Jordaan. Purification was conducted according to the manufacturers protocol; briefly, Ni^{2+} chelated magnetic beads were vortex mixed and 100 µL of suspension transferred to a new tube. Collected beads were washed twice and then equilibrated in binding buffer (80 mM sodium phosphate, 40 mM imidazole, 1.0 M NaCl, pH 7.4) for 5 min, collected via magnetic pelleting and excess buffer discarded. Previously collected soluble fractions (300 µL) were then introduced to the beads and topped up with an additional 400 µL of binding buffer and vortex mixed thoroughly. Binding was allowed to continue for 10 min at room temperature under agitation. Beads, with bound protein, were pelleted by magnetic collection and the supernatant discarded. Pellets were washed 5 times in binding buffer and supernatants from the wash steps discarded. Bound proteins were eluted by addition of 100 µL of elution buffer (80 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole pH 7.4) for 5 min at room temperature under agitation. Beads were pelleted and supernatant containing purified CNA35 probes collected (approx. 100 µg/workflow) and transferred to a new tube and stored at -20°C until use.

2.2.10.5. Preparation of type 1 collagen from rat tails

The use of collagen in bioengineering applications has been described extensively with relevance to bio-scaffolds, coatings and hydrogels [186-188]. Rat tails, used for collagen harvest, were kindly donated by Prof Michael Pepper from the Institute for Cellular and Molecular Medicine (ICMM) at the University of Pretoria. Collagen type

I fibres from rat tail tendons were isolated, collected and reconstituted using an adapted version of the protocol described by Rajan *et al.* [189]. Rat tails were washed in ultrapure water, dried and thin extremities twisted, and skin broken off. White collagen fibres were exposed, sheaths nicked using a surgical blade and fibres gently recovered using surgical tweezers and transferred to a beaker containing PBS. This process was repeated for the full length and both sides of the tails (n =12). Collected fibres were transferred to 100% acetone (5 min) and then to 70% ethanol (v/v) for a further 15 min. Fibres were then solubilized in a covered beaker containing 20 mM acetic acid with constant stirring (4°C; 48 h). Resulting opaque collagen containing solution was transferred to new tubes and centrifuged to remove insoluble components and impurities (10 000 *g;* 30 min) and supernatant collected and stored at -20°C until use.

2.2.10.6. Coating of coverslips for collagen probe validation

Type I rat tail derived collagen was thawed at 4°C, aliquoted, diluted and quantified using the previously described BCA assay. Collagen, for coating, was neutralized (pH 6-7) using 100 mM sodium hydroxide solution in water and diluted in ice cold PBS (pH 10-11). The required collagen concentration was calculated using the formula:

$$\text{Collagen Conc [µg/mL]} = \frac{\text{A}_{\text{Coating}}\,[\text{cm}^2] \times 5\,[\mu\text{g/cm}^2]}{\text{V[ml]}}$$

Collagen solution was added to 12 mm cover slips within 12-well plates, neutralized, and incubated over night at 4°C. Plates were then warmed in an incubator at 37°C for 2 h. After coating, excess collagen solution was aspirated off and coverslips dried before washing 3 times with PBS, blocking in 1.5 mL of 1% BSA in PBS and washing again 3 times with PBS. Coverslips were incubated with 1 µM of each CNA35 probe for 3 h at 37°C before rinsing 3 times with PBS, fixing with 4% PFA and mounting to a microscope slide. Binding of CNA35 probes were visualized using a Zeiss LSM800 confocal laser scanning microscope as described in [182].

2.2.10.7. Generation of protein-mass profiles from HepG2 cultures

Proteins recovered from HepG2 cell monolayers and spheroids were separated using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Briefly, 10 μg of protein was combined 3:1 with 4x Laemmli sample buffer 0.250 M Tris-HCL (pH 6.7), 8% SDS, 40% (v/v) glycerol, 20% (v/v) β-mercaptoethanol, 0.004% bromophenol blue) then incubated at 95°C on a dry heat block for 5 min, cooled and pulse centrifuged before loading onto Novex precast 4-20% gradient polyacrylamide Tris-Glycine gels (Invitrogen, Waltham, Massachusetts, USA). Precision Plus All Blue (BioRad, Hercules, USA) protein standards were used as mass marker. Proteins were separated in a Mini-PROTEAN Tetra System in running buffer (0.1% SDS, 25 mM Tris-base, and 19.2 mM glycine) at 80 V for 15 min then 160 V until the buffer front reached the end of the gel. Gels were fixed in 1% acetic acid (v/v) with 50% methanol (v/v) in dH2O for at least 30 min then stained with Coomassie brilliant blue (0.1% Coomassie R250, 5% acetic acid (v/v) with 50% methanol (v/v) in dH2O). Gels were then washed with dH2O multiple times on an orbital shaker until background was reduced and protein bands clearly visible. Destained gels were visualized using a ChemiDoc™ Gel Imaging System (BioRad, Hercules, USA) and analysed with Image Lab software (BioRad, Hercules, USA).

2.3. Results and discussion

2.1.1 Culture and characterization of HepG2 monolayer and spheroid models

Key benefits of monolayer cultures lie in their ease of use and being the historically standardized application. In the case of hepatocytes, this method of culture often oversimplifies the culture environment leading to altered gene expression that impacts the cells phenotype. Generally, cells that are immortalized possess inherent chromosomal abnormalities and mutational profiles enabling them to replicate indefinitely [190]. Prolonged proliferation enables further genetic drift and resultant clonal selection of a unique, highly proliferative cell clone over time. Primary hepatocyte cultures do not remain viable in monolayer conditions for long enough to

mimic this phenotypic drift process, however these cultures demonstrate a different, more immediate loss of cell specific processes and relatively rapid functional de-differentiation [191]. The degree to which immortalized hepatocytes can maintain their functional representation, while preserving replicative immortality underpins the degree of their usefulness as preclinical models.

Cell morphology and function are closely related characteristics. HepG2 cells (Figure 2.4 A-C) display an epithelial like morphology in contrast to the typical cubic cell shape of PHH (Figure 2.4 D) adapted from [192]. PHH also show varying subsets of binucleated cell populations, a feature largely absent from HepG2 cells. However, similar to PHH, HepG2 cells contain large vacuoles with cytoplasmic lipid droplets (Figure 2.4 C) and both cell types are capable of developing primitive bile canaliculi [193-195].

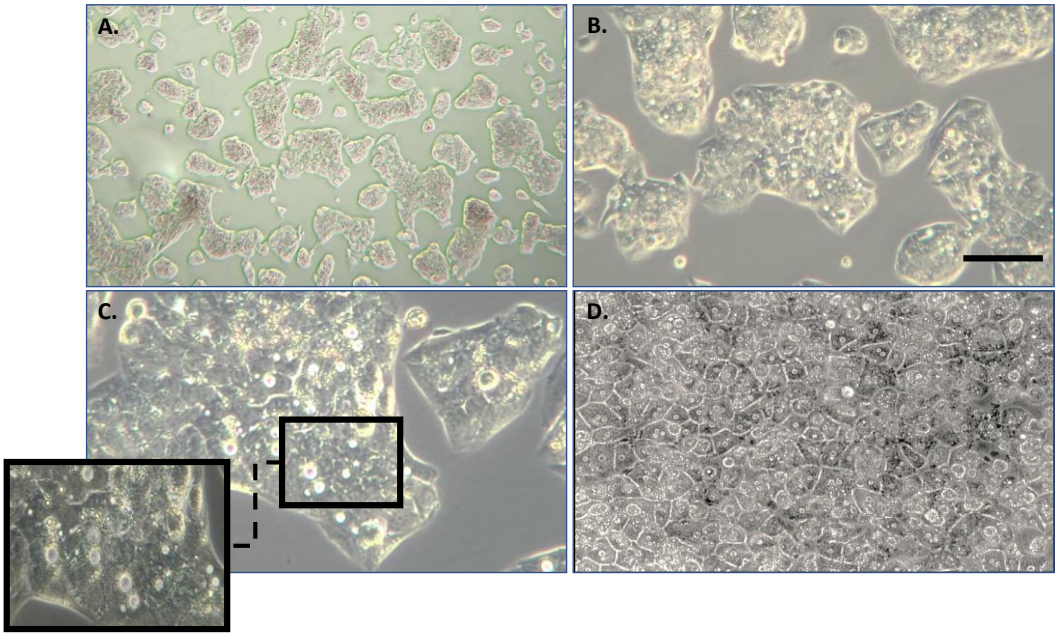


Figure 2.4: Phase contrast images of HepG2 cells (A-C) and PHH (D) cells cultured as monolayers. Scale bars: 50 µm. PHH image from [192]

The ability for preclinical models to accurately mimic the *in vivo* scenario has long been pursued yet often misrepresented. The idea that a collection of homogenous cells plated onto a dish surface can recapitulate the function of complex and highly heterogenous organ systems is unrealistic. Through understanding the shortfalls of a

cell culture model system, researchers have been able to adapt the conditions for these models to become slightly more representative. An early example of these adaptations has been the use of 3D modelling. PHH cultured in bioreactors express hepatocyte-specific functional proteins [196], and re-establish expression of hepato-specific metabolic genes (AhR, CAR and PXR) to the same or higher levels than what is observed *in vivo*. The selection of 3D culture method notably impacts the degree to which these changes occur, with bioreactor cultures better than hanging drop cultures with regards to growth parameters. Similarly, when HepG2 cells are cultured as spheroids using different methods, variable growth characteristics are observed.

Factors that influence the growth kinetics of monolayer cell cultures, and, by extension, spheroid cultures are nutrient accessibility, growth factor binding, chemical gradients, environmental pH, nature of substrates and the cell's genetics. In Figure 2.5, spheroids grown in hanging drop format were seeded at 20 000 cells per well compared to those seeded at 1000 cells per well in the 3D micro-moulds. While morphological features were well preserved, total protein content suggests important differences in cell growth. Spheroids in the hanging drop model, starting out with a higher seeding density, appear to proliferate quickly up to Day 14, after which minor increases in cell number are observed between Day 14 and 28 (Figure 2.5 C). However, seeding initialized with a lower cell number in the micro-moulds, showed a sustained relative protein content increases up to Day 28. These results are similar to those observed by Stampar [197] where HepG2 spheroid cross-sectional areas were calculated (Figure 2.6).

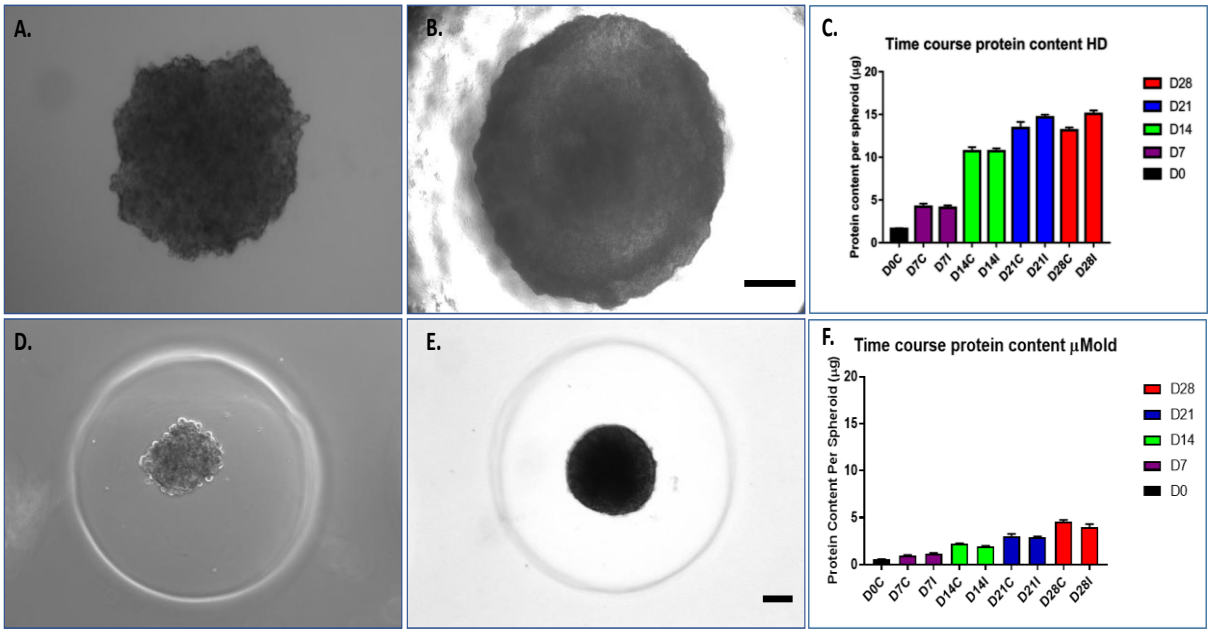


Figure 2.5: Micrographs illustrating spheroid morphology at Days 3 and 14 in culture. Bar graphs showing changes in protein content over the 28-day time course for the hanging drop model (HD) (A-C) and micro-mold µMold (D-F). cultures. D0C stands for the initial protein content of the cells used for seeding. Increases in protein content stands as an inference for cell proliferation across methods. Scale bars: 100 µm

Spheroids are formed through the active attachment and re-configuration of single cells to form higher order structures. This process is accompanied by changes of the cellular and extracellular organization. Increased cell-cell and cell-matrix interactions induce modifications to cytoskeletal structures which trigger differential regulation of their associated signalling networks. Activated cadherins transduce signals through their intracellular domains by interacting with the adhesion complex, which includes various catenin subunits and other cytoskeletal molecules. These complexes are able to trigger several other significant signalling cascades like β-catenin, a downstream regulator of the Wnt pathway that among others regulates cell fate, proliferation and migration [198, 199]. Adhesion molecules also influence growth kinetics through mechanical forces which together with a shift to cadherin-mediated binding have been shown to reduce cellular proliferation through a process known as contact mediated inhibition [200]. These molecular changes, together with the formation of differential zones within spheroids (Figure 2.7), may be responsible for the increase in phenotypic relevance of spheroid models.

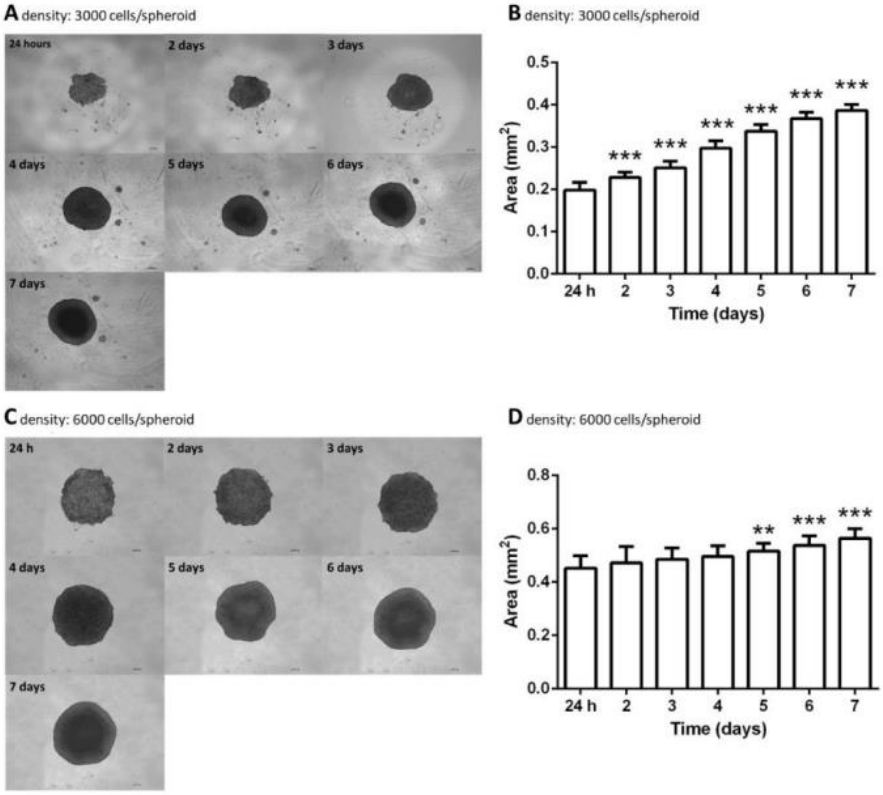


*Figure 2.6: The growth and morphology of spheroids (planimetry) monitored over 7 days of cultivation. The surface area of spheroid size was measured every 24 h (A–B: initial density of 3000 cells/spheroid and C–D: initial density of 6000 cells/spheroid). The images were taken using an inverted microscope at 40× magnification (N = 3). Results are presented as the mean ± SD (N = 10). The statistical analysis was performed in GraphPad Prism 6, by the one-way ANOVA using the Dunnett's multiple comparisons tests, ** p < 0.01, *** p < 0.001. Reprinted with permission under the Creative Common CC BY license. Figure taken from [197]*

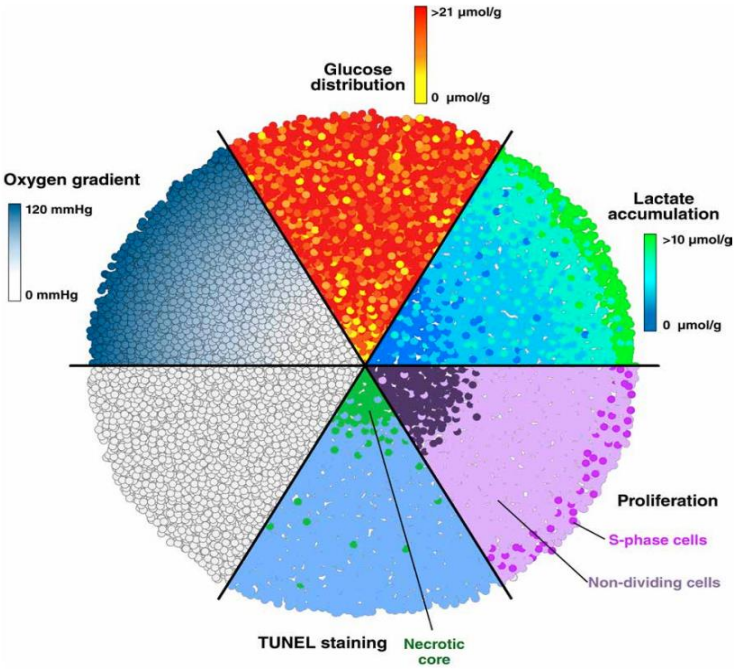


Figure 2.7: Zones in a spheroid tumour culture. Oxygen and nutrients fail to permeate to the necrotic core of the spheroid, and proliferating cells are more likely to be seen in the outer perimeter of the spheroid than within. Reprinted with permission from [201]

2.1.2 Assessing cytotoxicity in response to drug cocktail

Despite their reported predictive insufficiencies, HepG2 cells have been used extensively to identify potentially hepatotoxic drugs. For the purpose of long-term drug exposure studies absence of drug associated cytotoxic response must be confirmed to ensure the cells in culture remain viable. Assessment of acute innate cytotoxicity is easily inferred from a simple cell density assay such as the SRB cell enumeration assay as illustrated in Figure 2.8. However, this assessment method does not easily monitor long time course exposure, and even less so in the case of 3D cultures such as spheroids. It is therefore useful to monitor other cytotoxicity end points, such as LDH release, which can be adapted to assay spheroid models. Hanging drop based HepG2 spheroids were cultured for 28 days in the presence of a seven-drug cocktail at 5 µM for each drug to ensure absence of cytotoxic effects.

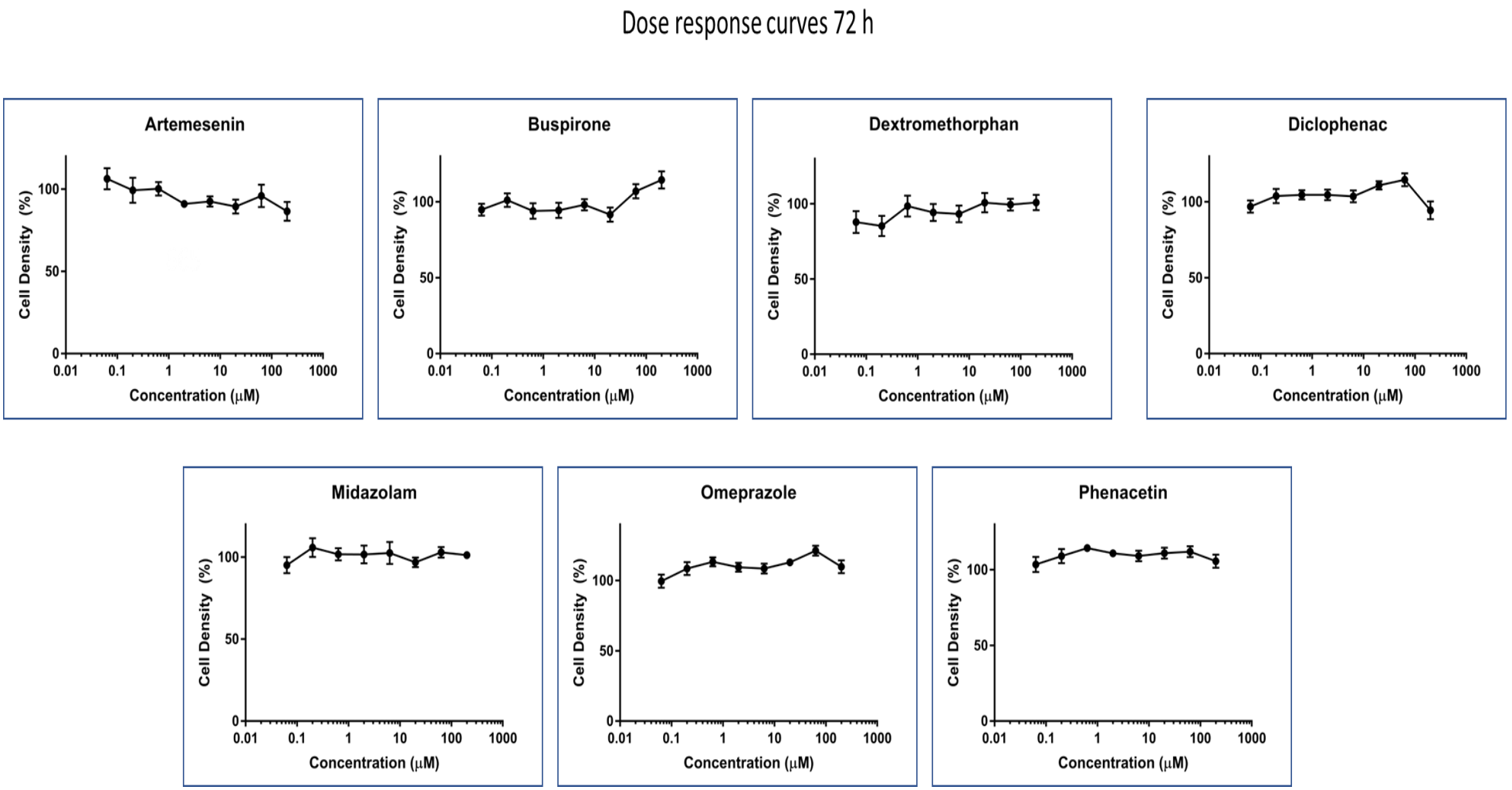


Figure 2.8: Dose response curves representing relative cell density in monolayer cultures of HepG2 cells as measured by the SRB assay

When adapting assays outside of their intended use, it is important to consider which assay parameters may be compromised and how to account for these. The LDH assay infers cytotoxicity from the release of LDH by membrane compromised cells into the media. With respect to its use in monolayer cultures, the released LDH accumulates over the course of exposure such that a terminal sample which is representative of the total LDH is assessed. In the case of spheroids, that are maintained for long culture times, regular media exchanges are necessary, which results in removal of the accumulated LDH. To account for this, control spheroids were cultured in parallel to the drug exposed spheroid group and monitored at the same time points. Potential differences in size were accounted for by normalisation to total protein for each respective group. Lack of significant cytotoxicity was observed between control groups and those exposed to 5 μM of drug cocktail (Figure 2.9) ensuring the viability of downstream proteomic assessments which took place at much lower drug exposure concentrations.

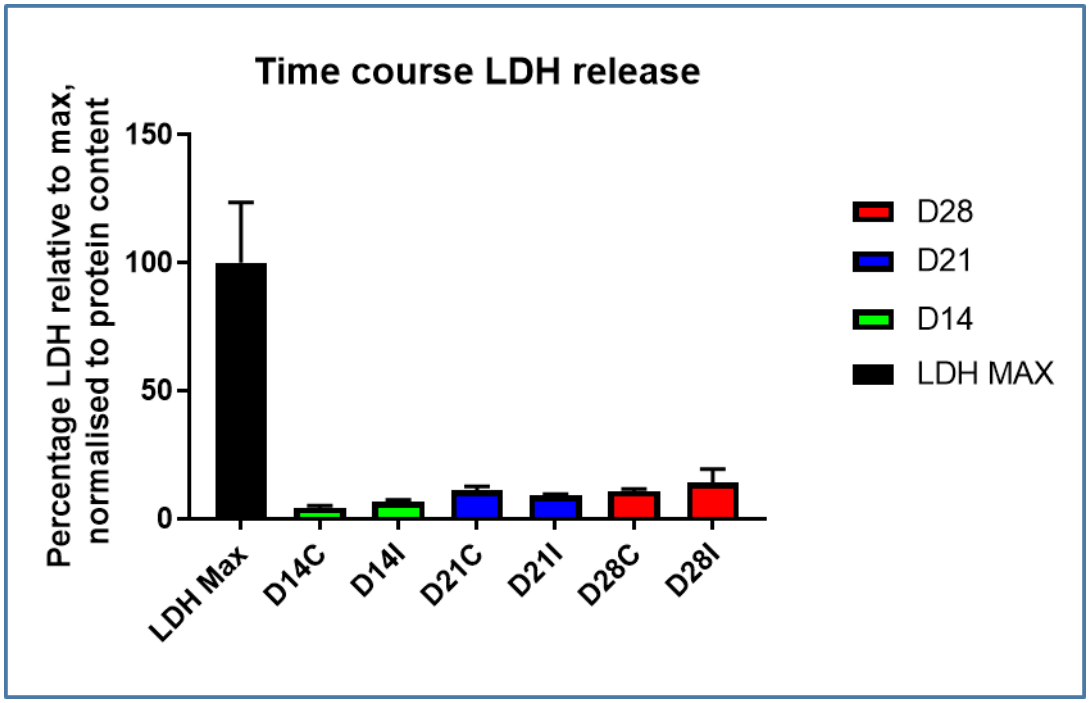


Figure 2.9: Time course LDH release for spheroid groups at Day 14, 21 and 28 (D14, 21 and 28, respectively). The C denotes control spheroids and the I denotes drug-exposed spheroids. Cytotoxicity is expressed as percentage relative to the LDH max control after normalisation to the protein content.

To further characterize and confirm the capacity for long-term culture viability, spheroids were assessed using the Live-Dead assay. Control spheroids at Day 28 in culture were exposed to 5 and 10 µM puromycin for 16 h before assessment. Puromycin exerts its toxic effect through inhibition of protein synthesis by ribosome-catalysed incorporation into the C-terminus of elongating nascent chains, blocking further extension and resulting in premature termination of translation [202]. Acute puromycin cytotoxicity (Figure 2.10) was observed in spheroid cultures with the response being proportional to dose for both light micrographs (top row) and Live-Dead stained confocal micrographs (bottom row). Control spheroids exhibited well defined borders and high FDA staining with low PI staining. Puromycin exposed spheroids in contrast, showed progressive breakdown of spheroid edges, increased PI staining and structural enlargement indicative of cytotoxicity.

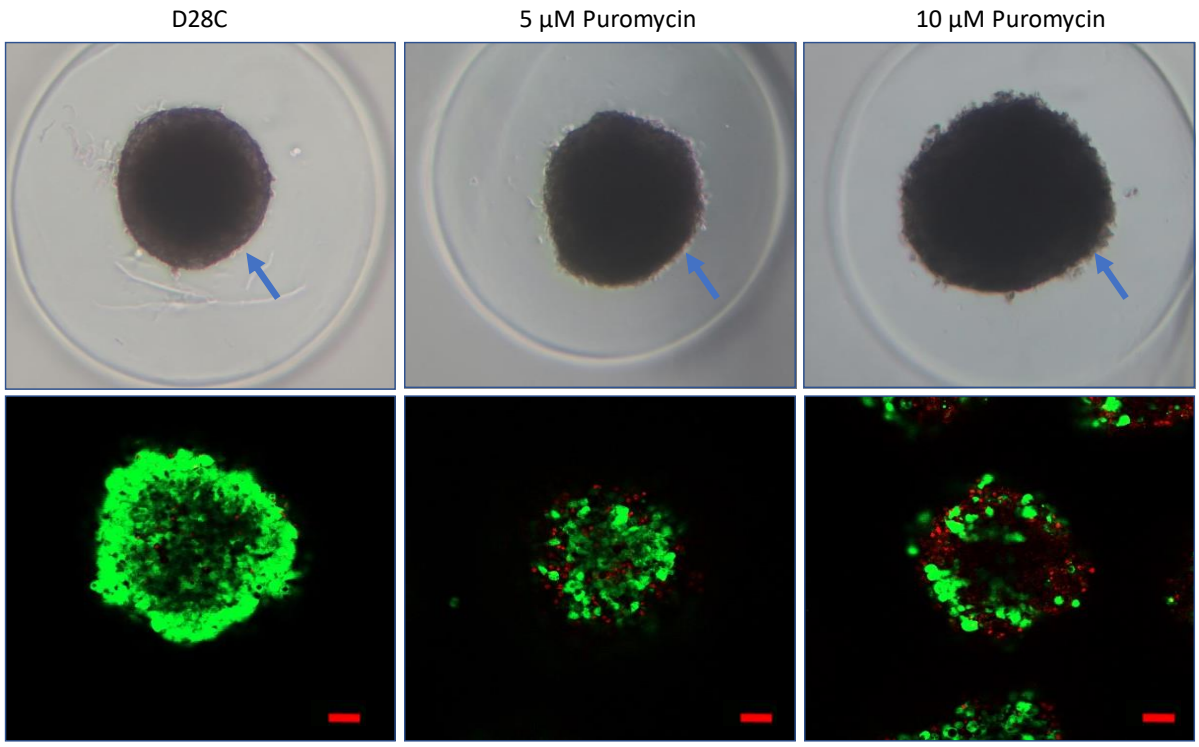


Figure 2.10: Light and confocal micrographs of Day 28 cultures of control (left), 5 µM (middle), and 10 µM (right) puromycin treated spheroids showing the progressive breakdown of spheroid edges and increasing PI staining indicating cell membrane damage. Scale bars: 100 µm

2.1.3 Structural changes as a result of spheroid cultures

Spheroids express a higher abundance and more complex network of extracellular matrix proteins compared to monolayer cultures. The basement membrane of healthy human livers is made up of appropriate ratios of glycoproteins, collagens, and proteoglycans. Collagen forms the primary structural component of the ECM which helps to regulate cell adhesion, chemotaxis, and tensile strength. Collagen deposition is important for spheroid growth and has been used as a supplemental lattice to facilitate robust cell growth. CNA35 is a protein expressed in bacteria which has a pan-collagen binding affinity. This protein has been functionalised with fluorescent markers making it useful for non-invasive collagen imaging [203]. Based on this ability, expression and purification of CNA35 was carried out to be able to visualize collagen deposition in different cell cultures.

CNA35 was expressed with either mTurquoise or tdTomato fluorescent labels and validated for binding using rat-tail collagen coated coverslips (Figure 2.11). However, due to its emission wavelength overlapping with that of many common nuclear stains, the mTurquoise labelled CNA35 was not used in subsequent assays. HepG2 cell collagen deposition was compared between both control monolayer and spheroid cultures over the culturing time course (Figure 2.12) where collagen expression was observed in both monolayer and spheroid conformations. However, collagen abundance was more pronounced in spheroid cultures with a progressive increase in abundance based on culture time.

Another useful marker protein for hepatocytes is Cytokeratin 18 (CK18), which is part of a major intermediate filament protein in the liver. It has a cytoprotective function, is expressed in both major hepatocellular types, hepatocytes and cholangiocytes, and is used as a diagnostic marker of hepatic injury or cell death. The clinical measurement of CK18 is used to identify the mechanism of hepatocellular death, as stable CK18 fragments indicate apoptosis due to CK18 being a target for caspase mediated proteolysis while the release of full length CK18 is indicative of necrosis [204-206]. Both forms of CK18 can be measured using ELISA clinical diagnostic tests. Figure 2.13 indicates the progressive upregulation of CK18 over the spheroid culture time course with a progression pattern similar to that observed for the upregulation of collagen.

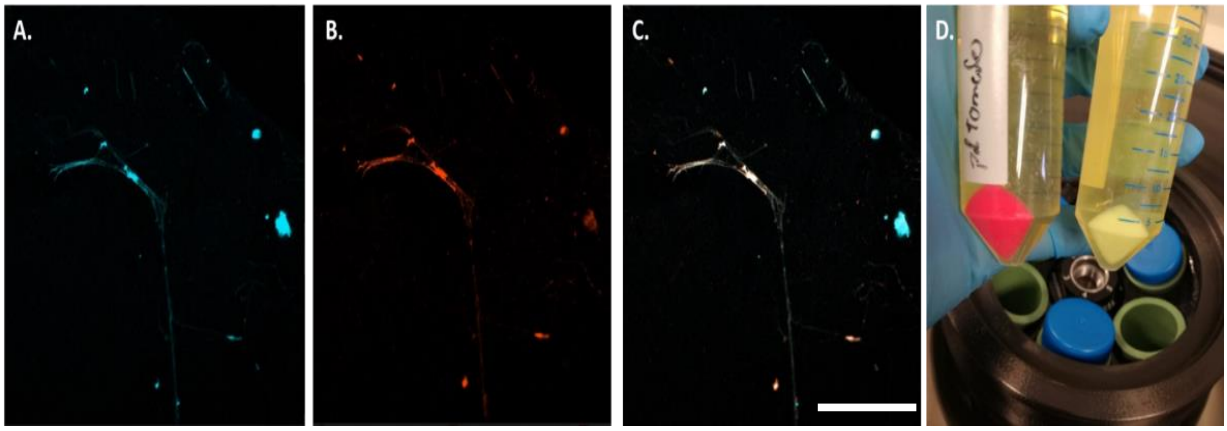


Figure 2.11: CNA35 collagen probe demonstrating binding to collagen coated plate. A) CNA35 mTurquoise, B) tdTomato, C) Merge of A and B representing areas of overlap, and D) pellets of transformed bacterial colonies induced using IPTG. Scale bar: 20 µm

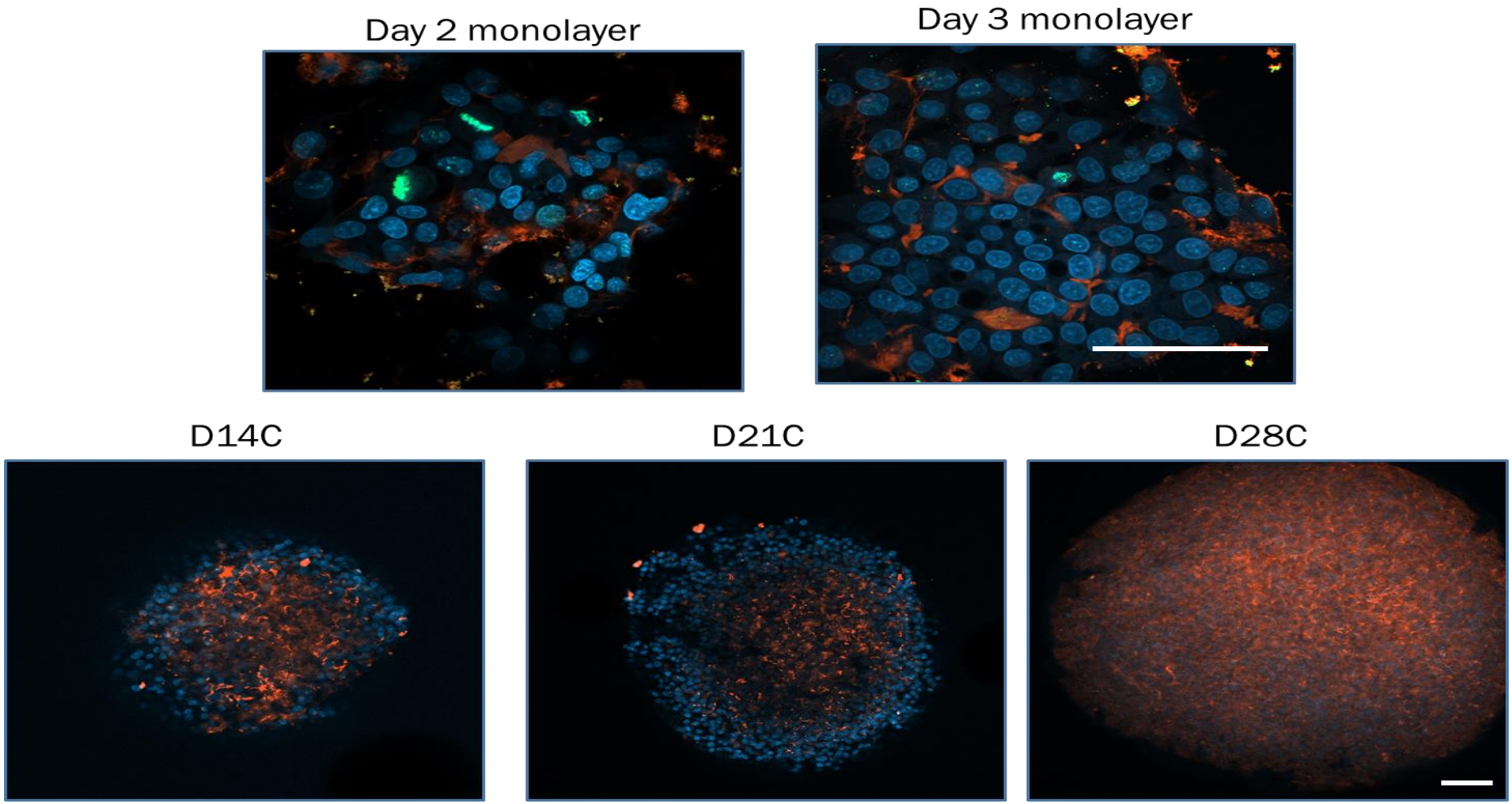


Figure 2.12: Pan-collagen binding by the CNA35 collagen probe (orange) in monolayer cells at Days 2 and 3 in culture (top row) and spheroids at Days 14, 21 and 28 in culture (bottom row). Scale bars: 100 μm

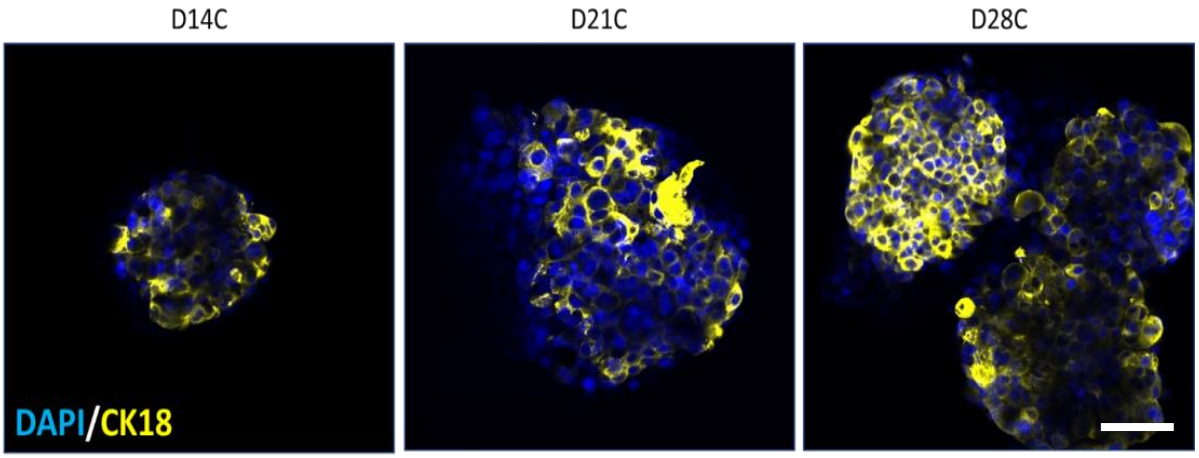


Figure 2.13: Confocal micrographs of CK18 (yellow) and DAPI nuclear stained (blue) spheroids from Days 14, 21 and 28 in culture. Scale bar: 100 µm

2.3.1. Protein-mass profiles of HepG2 spheroid and monolayer cultures

Protein mass profiles were generated using proteins collected from HepG2 monolayers and spheroid cultures to assess protein extraction and to qualitatively visualise protein expression patterns across culture conditions. Figure 2.14 shows a representative image for proteins extracted from various spheroid timepoints compared to monolayer.

It is important to note that a single band on a gel is generally made up of multiple proteins despite this, some differences in the mass ranges of 25-75 kDa are observed when comparing spheroid cultures (lanes 2-4) to monolayer (lane 5). Which may be indicative of protein expression differences between these culture methods.

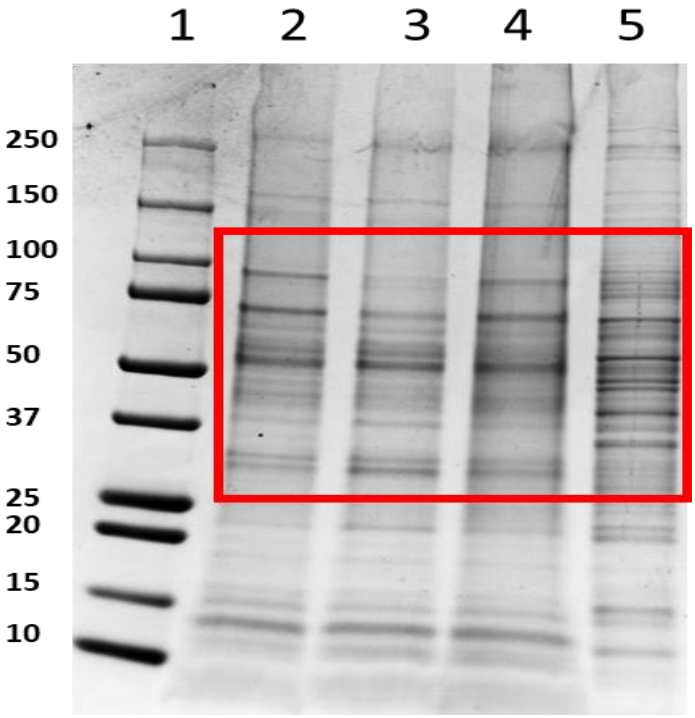


Figure 2.14: Coomassie stained Novex 4-20% precast gel of Precision Plus All Blue protein standards 10-250 kDa (1), HepG2 spheroids at days 14 (2), 21 (3) and 28 (4) in culture and HepG2 monolayer culture (5) harvested at ~80% confluence. Red box indicative of regions of proteomic change.

2.2 Conclusions

Spheroids were characterized with respect to a number of features that have potentially important outcomes for hepatic models in the *in vitro* setting for answering biological questions. HepG2 cells, cultured as spheroids, display a wider array of structural and biological variances compared to their monolayer counterparts, including the time that they can be maintained in culture presumable due to reduced growth kinetics conferred in proportion to time spent in culture. Hanging drop models, while beginning with higher seeding densities, reach a growth plateau after approximately 21 days in culture, a feature not observed in the equivalent culture time frame for micro mould cultures.

HepG2 cells did not display evidence of cytotoxicity in response to the drug exposures in the case of dose response curves generated for monolayer cultures and the combined drug cocktail at 5 μM exposed to spheroids, and this holds true even for repeated

exposures over an extended period of time. Additionally, qualitative changes to the proteomes of HepG2 cells between monolayer and spheroid cultures are evident as seen within the protein mass profiles generated during SDS-PAGE analysis. These data serve as preliminary evidence that that 3D cultures should express proteomic signatures which differ from their 2D counterparts. More thorough investigation of the proteomic changes to the HepG2 cell proteome, conferred by means of culture as spheroids and the further perturbations resulting from long term drug exposure will be discussed more thoroughly in subsequent chapters.

Article

Chapter 3. Continual proteomic divergence of HepG2 cells as a consequence of long-term spheroid culture

Andrea Antonio Ellero[1, 2] (u13209656@tuks.co.za), Iman van den Bout[2, 3] (Iman.vandenbout@up.ac.za), Mare Vlok[4] (marevlok@sun.ac.za), Allan Duncan Cromarty[1] (duncan.cromarty@up.ac.za), and Tracey Hurrell[5*] (THurrell@csir.co.za)

[1] Department of Pharmacology, Faculty of Health Sciences, School of Medicine, University of Pretoria, South Africa

[2] Centre for Neuroendocrinology, Faculty of Health Sciences, School of Medicine, University of Pretoria, South Africa

[3] Department of Physiology, Faculty of Health Sciences, School of Medicine, University of Pretoria, South Africa

[4] Proteomics Unit, Central Analytical Facility, Stellenbosch University, Stellenbosch, South Africa

[5] Bioengineering and Integrated Genomics Group, Next Generation Health Cluster, Council for Scientific and Industrial Research, Pretoria, South Africa

* Correspondence: THurrell@csir.co.za ; Tel.: +2712 841 4507

Link to article: https://www.nature.com/articles/s41598-021-89907-9

Link to supplementary information: https://www.nature.com/articles/s41598-021-89907-9#Sec16

Data availability: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD024353 and https://doi.org/10.6019/PXD024353.

Abstract: Three-dimensional models are considered a powerful tool for improving the concordance between *in vitro* and *in vivo* phenotypes. However, the duration of spheroid culture may influence the degree of correlation between these counterparts. When using immortalised cell lines as model systems, the assumption for consistency and reproducibility is often made without adequate characterization or validation. It is therefore essential to define the biology of each spheroid model by investigating the proteomic dynamics, which arise due to culture duration. As an example, we assessed the influence of culture duration on the relative proteome abundance of HepG2 cells cultured as spheroids, which are routinely used to model aspects of the liver. Quantitative proteomic profiling of whole cell lysates labelled with tandem-mass tags was conducted using liquid chromatography-tandem mass spectrometry (LC-MS/MS). In excess of 4800 proteins were confidently identified, which were shared across three consecutive time points over 28 days. The HepG2 spheroid proteome was divergent from the monolayer proteome after 14 days in culture and continued to change over the successive culture time points. Proteins representing the recognised core hepatic proteome, cell junction, extracellular matrix, and cell adhesion proteins were found to be continually modulated.
Keywords: Spheroids, dynamic proteome, stable isotope labelled proteomics, HepG2 cells

3.1. Introduction

The pharmaceutical industry implements increasingly stringent standards in drug development, and enforces persistent regulatory reviews of both new drug candidates (NDCs) and commercially available drugs [207, 208]. Only a small percentage of NDCs entering development fulfil the requirements to enter into Phase I clinical trials with, less than 10% becoming licenced products [209]. Failure of both candidates in development and marketed drugs impose great costs to both industries and consumers. Though numerous factors contribute to drug candidate failure or post marketed drug withdrawal, safety and efficacy together contribute 74-76% [210]. This is largely due to the limitations and inconsistencies of preclinical screening models. Exploring this, hepatotoxicity has been cited as the most common causative factor for withdrawal of post marketed drugs [211, 212]. When using immortalized or primary cell cultures as surrogates for their *in vivo* counterparts it is important to adequately characterize each model system before being able to make reliable biological inferences. Currently used liver models which include; primary human hepatocytes (PHH), transformed hepatocytes (HepG2, Huh7, and HepaRG), and hepatocyte-like cells (HLCs) derived from embryonic or induced pluripotent stem cells, are both genotypically and phenotypically distinct due to their various origins[213-215]. Each liver model has unique limitations, including the high cost and inter-donor variability of PHH [216], the lack of clinically relevant biotransformation capacity and strong cancer signatures of HepG2 cells [217], or the immature and variable hepatic phenotype associated with HLCs [218, 219]. These distinctive characteristics make each model fit for only certain applications. However, an advantage conferred to all *in vitro* liver model systems is the capacity for adaptation to three-dimensional (3D) culture, which supports the maintenance or acquisition of a more *in vivo* representative phenotype [19, 20]. This notion of improved

phenotype has been widely adopted by researchers, using cells of various origins, in attempts to recapitulate both organs and disease conditions such as tumours *in vitro* [22, 220, 221]. Duration of culture, presence or absence of extracellular matrix (ECM) proteins, static or perfused systems, ratio of cell number to culture media volume, and spheroid physical dimensions are all determinants of *in vitro* phenotypes [222], although a complete understanding of how each influence cell behaviour and phenotype remains unclear. Shifting the paradigm from monolayer to 3D culture systems has, in most models, not sufficiently addressed the progressive changes that occur as a consequence of duration of time spent in culture. Despite the widely accepted notion that cells cultured in 3D allow for more representable phenotypes, the rational for experimental time points is less thoroughly characterized. The impact of a dynamic baseline would be evident in experiments such as drug testing for efficacy or toxicity if cells which rapidly form spheroids were first cultured for 7, 10, or 28 days. While recognised as limited in their biotransformation capacity [223], HepG2 cell spheroids are used as a liver model to investigate genotoxicity [224], and predict hepatotoxicity [225, 226]. Phenotypic transitions as a result of growth in 3D have been investigated at the gene transcription level [227] which while informative, does not account for the fact that only 20-50% of transcribed genes result in functional expressed proteins [228], and could cast doubt on the perceived utility of model systems. HepG2 monolayers boast a highly reproducible detectible proteome while the proteome of HepG2 spheroids cultured for 10 days inconsistently diverges from monolayers [229] across this timeframe. Therefore, since proteomic changes which occur as cells arrange and mature within complex 3D structure are often overlooked [230], we sought to monitor the relative quantitative changes within spheroids over a long-term culture of 28 days using isobaric tagging. This was done to determine whether cells maintained in spheroids for more than 10 days undergo reproducible, significant, proteomic changes related to their altered culture architecture. These data suggest that there is a continual divergence of the HepG2 proteome as a

consequence of long-term spheroid culture which impacts cellular functionality. We provide new insights into the dynamic nature of 3D culture systems and highlight important considerations essential to the field of biological modelling.

3.2. Results

3.2.1. Spheroids undergo dynamic proteomic transitions based on time spent in culture

Successful isobaric tagging allowed confident identification and relative quantification of 4819 proteins shared across biological replicates within experimental groups (Supplementary Table 1). Hierarchal clustering (Figure 3.1a) segregated the monolayer proteome (D0) from all three spheroid groups D14C, D21C, and D28C. Changes in the relative abundances of protein clusters (blue, black, and red arrows) correlate with culture time. These trends demonstrate continual phenotypic evolution within the analysed spheroid cultures, which is reproducible across biological replicates, but dynamic within the observable temporal frame.

Distinct spatial clustering in PCA (Figure 3.1b) reinforces the observation that the culture method generates the greatest variance, with reduced but continual separation between spheroid cultures over time. Here, PC1 contributes 64.7% of the variance which appears to distinguish culture methodology. PC3 (Supplementary Figure 1), contributed only 4.4% of the variance, and was not able to spatially resolve groupings despite maintaining resolution of monolayers and spheroids by culture time along PC1. Contrary to observations in a previous study [229], the spheroid replicates produced reproducible proteome data when cultured for longer than 14 days. Correlation plots (Figure 3.1c) were poor when comparing monolayers to spheroids from each of Day 14, 21, and 28 replicates, illustrating the magnitude of proteomic divergence from monolayers. In contrast, high correlations between biological replicates of both monolayers and spheroids within the same experimental groups were observed. This high degree of correlation was seen to

diminish progressively with time, when comparing spheroid replicates from the earliest spheroid groups (day 14) to those from later groups (days 21 and 28).

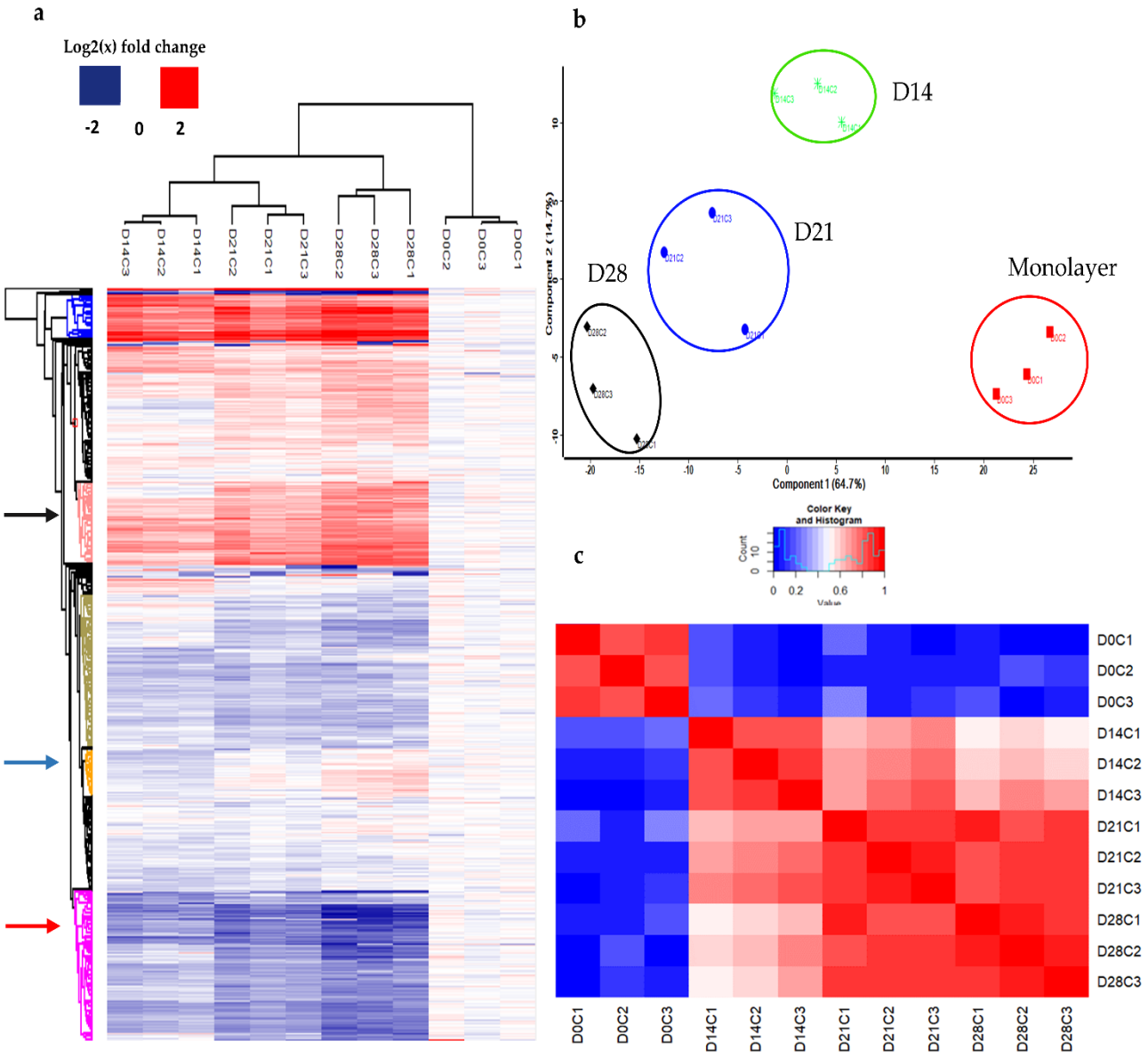


Figure 3.1. (a) Hierarchical clustering of proteomic data cohorts from biological replicates of monolayer (D0C1, 2, 3) and spheroid groups at Days 14 (D14C1, 2, 3), 21 (D21C1, 2, 3), and 28 (D28C1, 2, 3) in culture. (b) PCA of HepG2 cell monolayers (red) and spheroids at Days 14 (green), 21 (blue), and 28 (black) comparing Component 1 versus Component 2 and (c) Correlation plot showing strong correlations between biological replicates as well as moderate correlations across successive time points, whereas poor correlations were observed between D0C monolayers and spheroid cultures from D14C, D21C, and D28C. Figures generated using Perseus v. 1.6.7.0 software [231] and InfernoRDN v. 1.1.7626.35996 https://omics.pnl.gov/software/InfernoRDN [232].

3.2.2. *Direct comparisons demonstrate proteome differences resulting from culture methodology and duration*

To determine the number of differentially abundant proteins, volcano plots were generated using a two-sided t-test with an FDR of 0.05 while allowing for 250 randomisations and mean -log2(x) fold change of ± 0.3. Proteome data of spheroids from sequential weekly time points were compared to baseline monolayer cultures (D0) (Figure 3.2a, b, c; Supplementary Tables 2-4). Additionally, Day 14 (D14C) spheroid proteomes were compared with those of Day 28 (D28C) spheroids (Figure 3.2d; Supplementary Table 5). Using these significance thresholds, the proportion of differentially abundant proteins increased with spheroid culture duration. At Day 14, 38.72% of the detected proteome had superseded statistical thresholds compared to monolayers and this number increased along the temporal frame to 48.89% at Day 21 and 57.96% at Day 28. Furthermore, the changes between Day 14 and Day 28 involved 41.54% of the observed proteome indicating that the spheroid proteome continues to evolve over time.

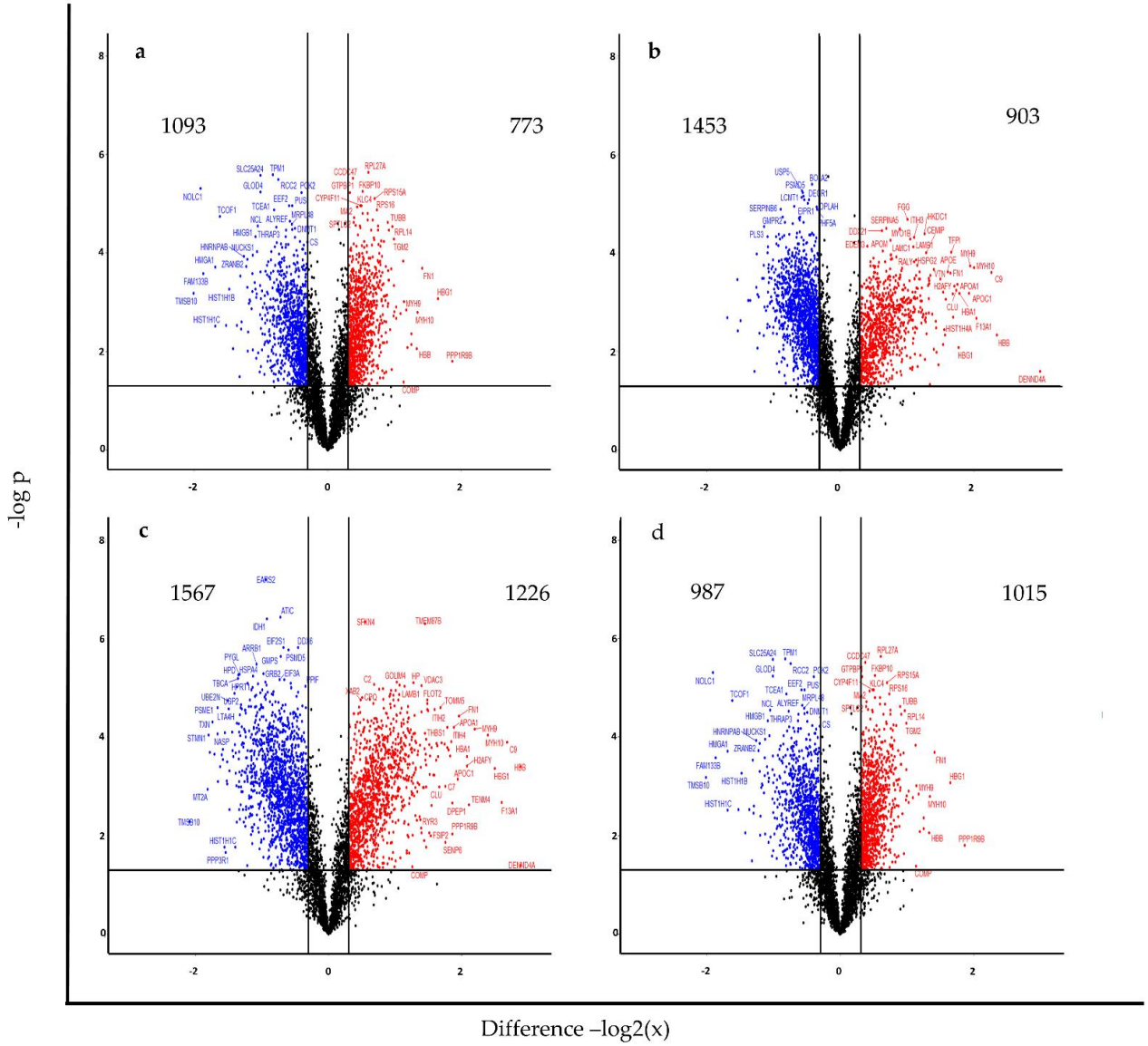


Figure 3.2 Volcano plots indicating significant differences in proteins across experimental groups identified using a two-sided t-test allowing for 250 randomisations, and FDR of 0.05 and mean fold change of log2(x) ± 0.3 for (a) D14C verses D0C, (b) D21C verses D0C, (c) D28C verses D0C, and (d) D28C verses D14C. Figure generated using RStudio software loaded with the ggplot package [233, 234].

3.3. Spheroids remain viable over a 28-day time course with a progressive increase in abundance of hepatic proteins relative to duration in culture

Protein content per spheroid (Supplementary Figure 3.2) was quantified throughout the time course as an inference for cellular biomass. Increases in protein content were greatest within the first two weeks of culture but then plateaued from days 14 to 28. Spheroids were monitored for morphological features including structural integrity and compactness (Figure 3.3a-d) as well as maintenance of esterase activity (Figure 3.3e-h, Supplementary Figure 3, Supplementary videos 1-2) as an indicator of cell viability. While the core proteomic assessments in this study were conducted over 28 days, these features were measured over a period of up to 6 weeks in order to demonstrate the long-term viability potential of this model system. HepG2 spheroids expressed hepatic marker proteins albumin (ALB), α-fetoprotein (AFP), and hepatic nuclear factor 4 alpha (HNF4α), as evidenced by confocal imaging (Figure 3.3i-p; Supplementary videos 3-5). Quantitation of AFP and ALB staining (Supplementary Fig 4) supported the trends in the proteomic data (Supplementary Table 1). HNF4α is a master regulator for hepatocyte differentiation [235] and is essential for the expression of other hepatic transcription factors and for maintaining hepatocyte function [236, 237], which remained consistently expressed throughout the time course. AFP, a glycoprotein derived from embryonic endoderm cells of foetal liver, decreases in abundance throughout hepatic maturation [238] whereafter ALB progressively increases in abundance in mature livers [239]. AFP and ALB increased in abundance in HepG2 spheroids, which given the cellular origin of these cells as being from a well-differentiated hepatocellular carcinoma, the progressive increase in AFP expression is expected [240].

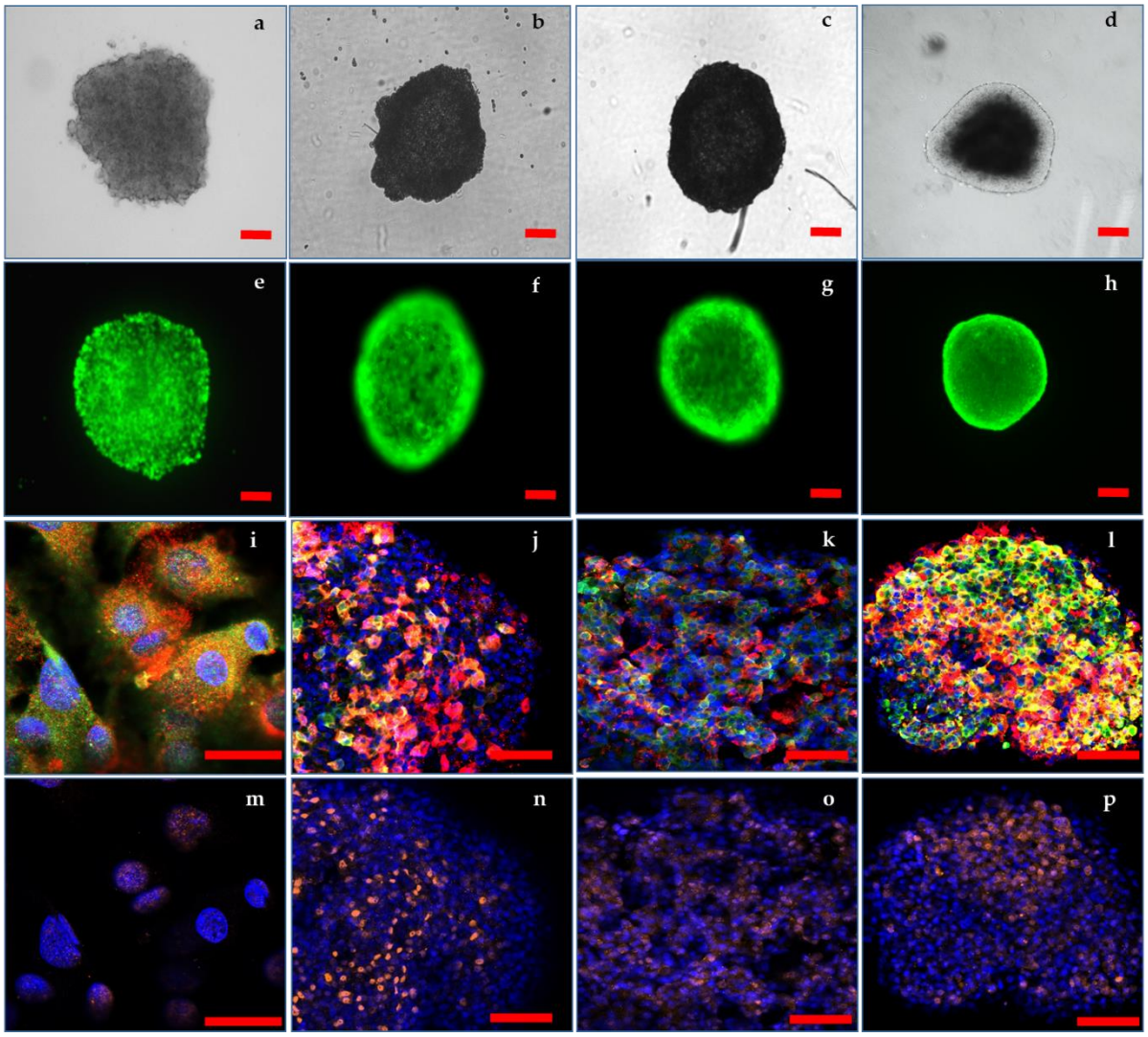


Figure 3.3: Images of the progression of HepG2 spheroids over a time course of 6 weeks in culture. Light micrographs depicting HepG2 spheroids at Days 14 (a), 21 (b), 28 (c), and 42 (d) post seeding. Fluorescent micrographs of HepG2 spheroids at Days 14 (e), 21 (f), 28 (g) and 42 (h) stained with FDA and PI showing maintenance of esterase activity and no evidence of compromised cell membranes. Confocal images of cleared HepG2 monolayers (i) and cleared spheroids at Days 14 (j), 21 (k), 28 (l), stained with DAPI (blue), ALB (Albumin, green) and AFP (α-fetoprotein, red). Confocal images of HepG2 monolayers (m) and cleared spheroids at Days 14 (n), 21 (o), 28 (p) stained with DAPI (blue) and HNF4α (orange). Scale bars 100 µm for spheroids and 20 µm for monolayers. Generated using Zeiss Zen Blue edition software 3.0 (https://www.zeiss.com/microscopy/int/products/microscope-software/zen.html).

The relative quantitative proteomic data generated allows for not only a global overview of the divergence of monolayer and spheroid cultures but also for interrogating the abundance of enriched protein data sets which underlie these differences. Spheroid

culture improves the expression of cell specific markers, suggesting that cells within spheroids become functionally closer to their *in vivo* counterparts [241-243], which has been observed in numerous hepatic cell spheroid cultures [218, 219, 227]. The potential for bias in the protein expression patterns within the enriched data sets due to the cellular origin of these spheroids could be a limitation. Therefore, datasets from this study were searched against a list of proteins which correlates with hepatic phenotypes from previous studies [218, 219, 244] and then enriched for analysis. Expression of proteins typically followed a progressive pattern whereby if a protein is up or down regulated at Day 14 this change continues and is amplified at Day 28 (Fig 3.4, Table 2; and Supplementary Table 6). Several proteins exhibited reduced abundance in spheroids when compared to the monolayer cells, namely; CEBPA (enhancer-binding protein alpha), EPHX2 (bifunctional epoxide hydrolase 2), MAT1A (adenosylmethionine synthase isoform type-1), CPS1 (carbamoyl-phosphate synthase), and AGXT (pyruvate aminotransferase). These proteins have a variable expression profile within hepatocellular carcinoma and liver tissue. Decreased AGXT expression is implicated in the progression of hepatocellular carcinomas [245], while CEBPA has a role in the regeneration of normal livers and typically shows down regulated levels of its mRNA expression during growth of freshly isolated hepatocytes [246]. This data shows that HepG2 cells express hepatocyte marker proteins more abundantly when cultured as spheroids for more than for 21 days, suggesting that they could attain a more hepatocyte-like phenotype than can be achieved during monolayer or early spheroid culture. Indeed, upregulated proteins fulfil a variety of hepato-specific functions such as those acting as nuclear factors (HNF4α and GATA4), apolipoproteins (APOA/B/C variants), and cytokeratins expressed in bile canaliculi of mature hepatocytes (KRT18). Their relatively increased abundance may still be insufficient, or conversely exceed what is observed within whole liver tissue but provides insight into the continual proteomic changes taking place during the time-course of spheroid culture. Cohorts were further enriched

for proteins involved in hepatic drug metabolism using the "core" and "extended" Absorption, Distribution, Metabolism and Excretion (ADME) lists (http://pharmaadme.org/; Supplementary Tables 7, 8) This ADME protein cohort was poorly represented in the data set, with only 94 proteins of the combined list of the 299 search proteins present. This could be due to the poor metabolic competence of these cells or alternatively due to a relative low abundance of these proteins within whole cell lysates. Where present, these proteins exhibited a variable expression change with time (Supplementary Figure 5).

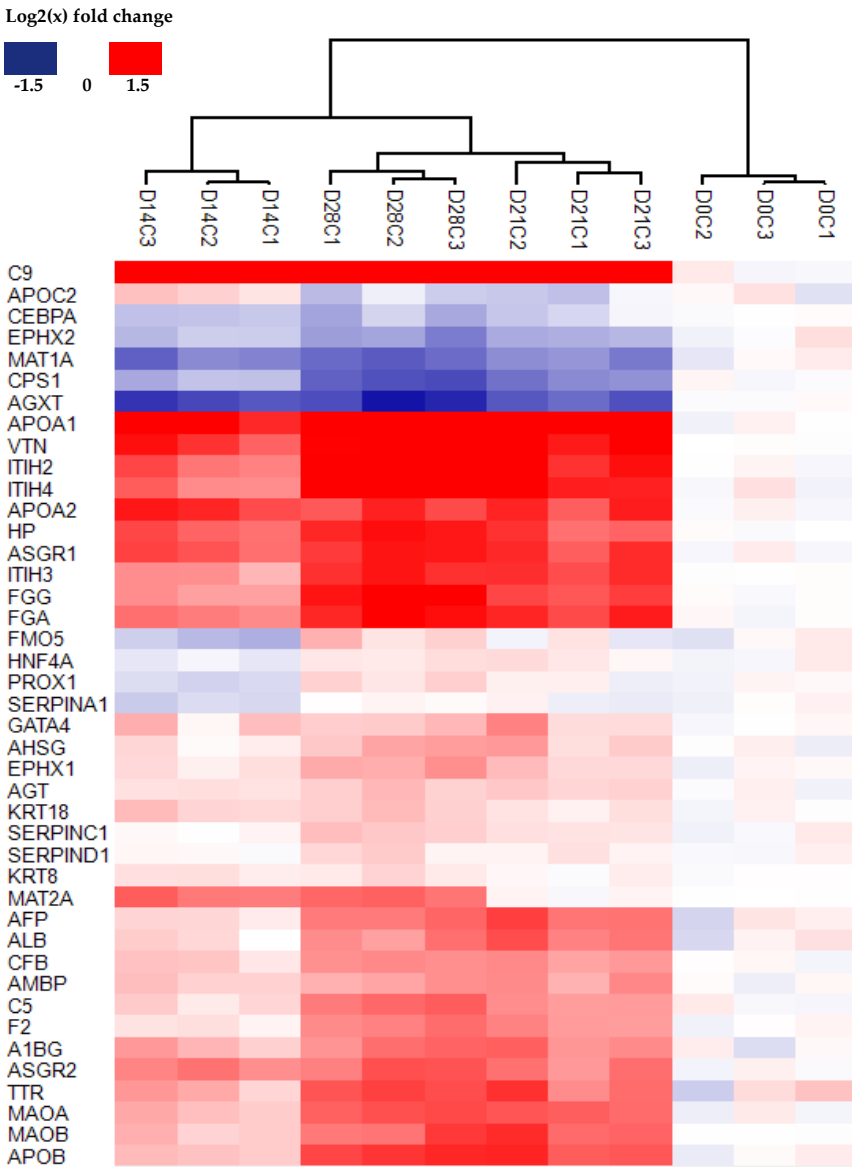


Figure 3.4: Hierarchical clustering of enriched proteins representative of a hepatic phenotype. Figure generated using Perseus v. 1.6.7.0 software [231].

Table 3.1: Expression of proteins representative of a hepatic phenotype as compared to monolayers.

Gene names	T-test Significant D14C	T-test Significant D21C	T-test Significant D28	Mean D14C	Mean D21C	Mean D28C
C9	+	+	+	1.91	2.27	2.69
APOC2				0.25	-0.24	-0.25
CEBPA	+		+	-0.33	-0.2	-0.41
EPHX2	+	+	+	-0.32	-0.44	-0.59
MAT1A	+	+	+	-0.74	-0.65	-0.85
CPS1	+	+	+	-0.39	-0.68	-0.94
AGXT	+	+	+	-1.02	-0.91	-1.16
APOA1	+	+	+	1.41	1.75	1.88
VTN	+	+	+	1.11	1.50	1.63
ITIH2	+	+	+	0.82	1.33	1.59
ITIH4	+	+	+	0.72	1.40	1.86
APOA2	+	+	+	1.16	1.12	1.05
HP	+	+	+	0.89	0.93	1.27
ASGR1	+	+	+	0.93	1.08	1.22
ITIH3	+	+	+	0.55	1.10	1.19
FGG	+	+	+	0.56	1.01	1.44
FGA	+	+	+	0.72	1.15	1.39
FMO5	+			-0.37	-0.02	0.28
HNF4A			+	-0.11	0.13	0.15
PROX1	+		+	-0.21	0.03	0.22
SERPINA1	+			-0.24	-0.04	0.03
GATA4			+	0.29	0.36	0.32
AHSG			+	0.12	0.35	0.45
EPHX1			+	0.16	0.26	0.51

AGT	+	+	+	0.17	0.26	0.30
KRT18	+		+	0.28	0.14	0.31
SERPINC1		+	+	0.03	0.16	0.31
SERPIND1				0.02	0.10	0.19
KRT8	+			0.15	0.04	0.16
MAT2A	+		+	0.79	0.03	0.82
AFP		+	+	0.19	0.87	0.77
ALB		+	+	0.17	0.81	0.65
CFB		+	+	0.27	0.58	0.63
AMBP	+	+	+	0.29	0.58	0.52
C5		+	+	0.21	0.57	0.82
F2		+	+	0.13	0.59	0.72
A1BG		+	+	0.41	0.70	0.75
ASGR2	+	+	+	0.69	0.72	0.88
TTR		+	+	0.43	0.86	0.99
MAOA	+	+	+	0.37	0.88	0.95
MAOB	+	+	+	0.32	0.95	0.86
APOB	+	+	+	0.33	1.01	1.11

[1] Values reported are the means, as Log2(x), of three biological replicate experiments showing relative fold change relative to monolayer counterparts. + indicates significance for q-value ≤ 0.05. Original non-averaged values may be found in Supplementary Table 1.

3.3.1. Curating biological processes to assess spheroid dynamics

In large proteomic data sets, biologically significant changes are not always of sufficient magnitude to reach statistical thresholds set prior to analysis. Spheroid cultures increase the proportion of cells in direct contact with each other. Intuitively, when considering the possible drivers of proteomic changes, the most likely contributors are the proteins which

facilitate these cell-cell interactions i.e. extracellular matrix (ECM) (GO:0031012), proteins which facilitate cell-cell attachment and regulate intercellular signalling pathways [247], which are associated with cell junction proteins (GO:0005911), and cell-cell adhesion molecules (GO:0098609). Therefore, the trends for proteins of these classes were investigated. To avoid missing potentially important divergent proteins or biasing based on statistical thresholds, protein annotation and functional classification was used, where

proteins were annotated and filtered based on the associated GO terms, and enriched using the PANTHER gene function classification system [248].

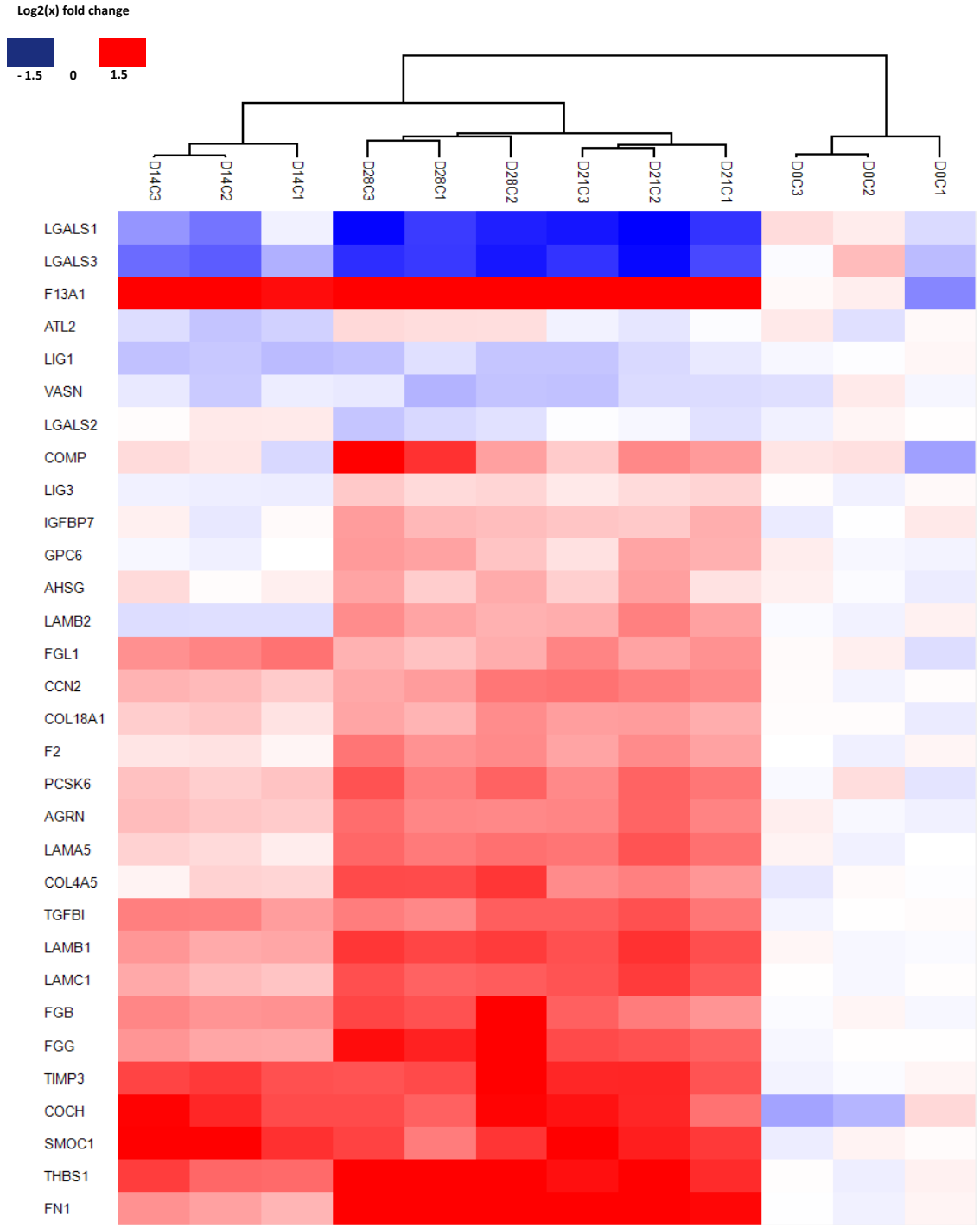


Figure 3.5: Hierarchical clustering of proteins involved in the ECM. Figure generated using Perseus v. 1.6.7.0 software [231].

Major ECM proteins (Figure 3.5; Supplementary Table 9) increased in abundance in spheroid cultures, seemingly increasing with time. The greatest fold changes were observed for various functional proteins produced by the liver; coagulation factor XIII (F13A1), tissue inhibitor of metalloproteinase 3 (TIMP3), and thrombospondin (THBS1) with notable changes also evident in laminin (LAMB1/2 and LAMC1) proteins. Galectin proteins 1 and 3 (LGALS1 and LGALS3) significantly decreased in expression over the time course contrary to what has been reported for hepatocellular carcinomas [249]. Interestingly, when enriching for proteins involved in focal adhesions (Supplementary Figure 3.6; Supplementary Table 10), integrin beta 1 (ITGB1) was found to be one among only 4 proteins of this class with a significant up regulation through the time course of spheroid culture. Previous evidence has shown that the expression of ITGB1 is required for the expression of LGALS1 [250], and despite the increased ITGB1 expression, LGALS1 was progressively decreased in spheroids. TIMP3 has been shown to have its expression silenced in hepatocellular carcinoma [251], contrary to the trend being observed in this study. Literature has reported that increased expression of specific laminin proteins is integral to the differentiation of stem cells into hepatocyte-like cells whereby progression of feto-hepatic phenotypes to a more mature hepatic phenotype is promoted [252].

Cell-cell junction (Supplementary Figure 7; Supplementary Table 11) and cell-cell adhesion proteins (Supplementary Figure 8; Supplementary Table 12) are involved in cell-substrate and intercellular attachment and regulate critical pathways such as monitoring of barrier functions in epithelia and playing critical roles in cell proliferation and cellular migration [253]. Integrins are the main cell-extracellular matrix adhesion molecules involved in the formation of focal adhesions and hemidesmosome junctions, while cadherins are usually involved in cell-cell adhesion molecules that forms adherens and desmosomal junctions. These anchoring junctions regulate the roles of intercellular molecules such as actin and intermediate filaments by dictating the cytoskeletal structure

of the cell. Cx32, a protein that forms gap junction channels for cell-cell communication [254], is progressively up-regulated throughout the time course. Though the abundance of proteins involved in cell-cell junctions and adhesions varied, where increased or decreased abundance was observed for a specific protein, this progression was again relative to duration in culture.

3.4. Discussion

Cell culture is integral in the development of therapeutics and continues to advance the understanding of biological pathways. The continual evolution of *in vitro* modelling has in recent times seen 3D cultures being widely adopted for new applications with several underlying assumptions being accepted. Spheroids allow for better biological mimicry of *in vivo* tissues when compared to monolayers [255], especially in the case of hepatocyte derived models [256, 257]. Furthermore, unlike monolayers which are generally limited in culture duration from seeding to confluence, spheroids may be cultured in excess of 5 weeks if given optimal conditions. This greatly extended culture time necessitates that the focus shift to whether a certain model is the most appropriate platform to answer a certain research question but also requires assessing at which point the model can most accurately and reproducibly do so? This study presents a model for HepG2 cell culture which has, within set statistical parameters, highlighted a complete divergence of the cellular proteomic of spheroid cultures from that of monolayers, contrary to what has been previously observed when assessing spheroids cultured for a duration of 10 days [229]. This divergence takes place reproducibly across biological replicates, within this system, if the spheroids were cultured for a minimum of 14 days from seeding. The multidimensional cell-to-cell interactions and improved paracrine signalling resulting from spheroid culture allows for new gene expression pathways to be activated and ultimately allows for a divergence in cellular phenotype in cells receiving these signals [258, 259]. The magnitude of this divergence is not well investigated, nor is it frequently

reported in literature. Cells within spheroids progressively deposit extracellular matrix thereby altering the abundance of proteins involved in cell junctions and cell adhesion [17]. The ECM is an intricate network comprising of multiple proteins, many of which provide important information for the building of a sophisticated structure required for anchoring cells and sustaining normal function of tissues. The matrix itself has been reported to be considered as a paracrine/endocrine entity, with more complex functions than previously appreciated [260]. Additionally, in the case of hepatocytes, the mechanisms by which these cells attach, changes from being largely integrin mediated in monolayer culture to being facilitated by cadherins and extracellular matrix proteins, which drives further phenotypic changes [261]. A separate consideration, not addressed here, is the variability with which cells from different origins form spheroids and how universal these time dependent proteomic changes are across different cell type models.

As reported before [229], spheroid proteomes do indeed diverge from monolayers, however, the time taken for this divergence to occur reproducibly across replicates was not consistent over the 10-day time course. Therefore, the culture period required to reach a stable or more biologically applicable phenotype is critical to the appropriate timing of experimentation using these models. Following this narrative, prior to acceptance as biological models, within academia or industry, cell phenotypes need to be thoroughly characterised. In this study it was observed that while spheroid cultures do converge in terms of proteomic phenotypes with time, the equilibrium reached is still not static but rather dynamic, and dependant on culture time. This makes standardisation challenging and researchers need to extensively characterise culture models with this consideration in mind. The focus for the development of 3D culture methodologies should be to develop reproducible and predictable 'fit–to-purpose' assays or models. While no single model currently exists to perfectly recapitulate the *in vivo* scenario in its appropriate

complexity, the use and development of appropriate and validated models remains paramount to basic biological sciences.

Although gene expression changes are typically several orders of magnitude higher than proteomic variances [262], the consensus and quality of the abundant protein cohort in spheroid cultures demonstrates that large scale proteomic changes should be used to characterize and define spheroid dynamics and not only evidenced at the transcript level. However, defining these 'fit-to-purpose' models requires the consideration that large fold changes in protein levels are not always essential to bring about biological changes, and that what is significant in a statistical sense is not always significant in a biological sense which is difficult to discern within global proteomic datasets.

As presented here, cells of a single origin are capable of undergoing significant changes to the proteome even when keeping all external conditions consistent and introducing a simple modulation of the culture environment over time. The changes observed do not fully endorse the narrative that spheroid culture allow for a more representative model of the *in vivo* scenario. While these cells are derived from a hepatocellular carcinoma the changes observed do not fully fit with the notion that the phenotypes of these spheroids become more like a hepatocellular carcinoma. While some cancer signature proteins are indeed increased as in the case of AFP (Figure 3.3g, h, i, 4 and Table 2). Others such as TIMP3 (Figure 3.4), AGXT and LGALS1/3 (Figure 3.5) changed their expression in the favour of a non-cancerous hepatic phenotype. This was coupled to the general increase observed in the expression of hepatic marker proteins. Proteome changes of this nature, whether considering this platform to model either hepatotoxicity testing or hepatocellular carcinoma cytotoxicity, cannot be overlooked.

The time course presented here is limited to 28 days, and as noted from the observations presented in Figure 3.1, these changes remain dynamic, and it is unclear as to how long this progression can continue as well as how beneficial or detrimental it will be to the cellular functionality. It is also unclear as to whether a static equilibrium will ever be reached within this spheroid system. Additionally, dissociation by TrypLE may have biased the baseline of membrane bound proteins in monolayers as compared to spheroids. However, as the same monolayer served as the day 0 protein sample and seeded the spheroids, the changes in the proteome over days 14 to 28 successfully characterize the temporal dynamics of the spheroid proteome. Despite these uncertainties, spheroid models such as this now form a central component in the pharmaceutical industry and academia [263, 264] and the critical task remains to interrogate long term spheroid dynamics and validate these models as appropriate for the assays for which they are intended.

This study provides insights into the dynamic nature of 3D culture systems and highlights important considerations essential to the application of a biological model. We have previously reported that spheroids cultured for 10 days in hanging-drops have a less reproducible proteome, across biological replicates, when compared with monolayers. Here it was demonstrated that increased culture time allowed for a continual divergence of the proteome of spheroid cultured cell replicates from monolayer cultures. There is a continuous, reproducible, divergence of proteome along the temporal frame with regards to proteins associated with hepatic phenotype, cell junctions, extracellular matrix, and cell-adhesion molecules. Taken together, these data demonstrate how highly dynamic the proteome of spheroid cultures are and provides a resource for assessing proteomic changes based on cell culture modulation. Adequate understanding of time-dependent changes derived from modulation in cell culture methodology in spheroid

cultures is essential in reducing proteomic heterogeneity and may ultimately allow for these 'fit-to-purpose' models to become reproducible even across laboratories.

3.5. Materials and Methods

3.5.1. HepG2 cell culture and spheroid formation

Human hepatoma HepG2 cells were obtained from Cellonex (Johannesburg, RSA: CHG2-C) and cultured in Dulbecco's modified minimum essential medium (DMEM) supplemented with 10% foetal bovine serum (FBS), and 2 mM GlutaMAX. HepG2 cells were thawed and cultured to 80% confluence (5-6 days) for 2 passages prior to harvesting monolayer controls (D0) using TrypLE and seeding spheroids. Spheroids were generated by seeding 20 000 cells/well in 45 µl of medium into Perfecta3D 96-well hanging drop plates (3D Biomatrix; Michigan, USA) according to the manufacturer's protocol. Cells aggregated under gravity at the droplet apex to form a single spheroid per well and partial exchange of growth medium (12 µl) was conducted every alternate day. Spheroids were harvested for analysis at culture Days 14, 21, and 28.

3.5.2. Sample collection and protein quantitation

Cells from dissociated monolayers were harvested as the Day 0 control. Cells were washed with PBS and lysed using 100 µl lysis buffer (10 mM Tris-HCL (pH 8), 1 mM ethylene diaminetetraacetic acid, 0.5 mM ethylene glycol tetraacetic acid, 1% Triton X-100, 0.1% sodium dodecyl sulphate, 0.1% sodium deoxycholate, 140 mM sodium chloride) containing cOmplete protease inhibitor cocktail. Cell lysis was aided by ultrasonic disruption on ice (1500 W, 5 min), and an additional 15 min incubation on ice, after which the lysate was centrifuged at 16 000 *g* for 10 min at 4°C. Soluble protein was quantified using the bicinchoninic acid assay. Absorbance was determined using a Bio-Rad iMark microplate reader at 560 nm. Spheroids (n = 80) were collected, combined, and

homogenised in 300 µl lysis buffer using a Dounce homogeniser prior to protein quantitation as described above.

3.5.3. Protein preparation and isobaric tag labelling

Biological replicates (n = 3) of HepG2 monolayers (D0), HepG2 spheroids at Day 14 (D14), Day 21 (D21), Day 28 (D28) and a pooled group comprised of equivalent amounts of protein from each experimental group (Pool) were each labelled with one of the 6-plex tandem mass tags (TMT; Thermo Scientific; Maryland, USA). Seventy-five micrograms (75 µg) of protein was reduced at 37°C for 1 h using 10 mM dithiothreitol, then alkylated at room temperature with 25 mM iodoacetamide. Proteins were precipitated overnight at 4°C after addition of 10 volumes of 100% acetone, harvested by centrifugation at 16 000 *g* (20 min) and resuspended in 100 mM HEPES (pH 8.5). Samples were digested with sequence-grade modified trypsin (1:40) for 1 h at 37°C with a further overnight digestion at 37°C after adding additional trypsin (1:40). Each TMT tag was resuspended in 41 µl mass spectrometry-grade acetonitrile. Supernatants from clarified (20 min; 16 000 *g*) digested peptides, were labelled for 2 h at room temperature under constant agitation. Labelling was quenched by adding 8 µl of 5% hydroxylamine for 1 h and incubated overnight at 4°C with dH_2O [265]. To circumvent tag affinity bias, experimental group's biological replicates were divided across tag sets and randomized across the 6-plex tag sets. The different tag labelled samples were combined to contain all corresponding 6-plex tags and dried under vacuum centrifugation.

3.5.4. Solid phase extraction and peptide fractionation

The combined labelled peptides were solubilized in dH_2O with 0.1% trifluoroacetic acid and loaded onto a conditioned SepPak C18 cartridge (100 mg), desalted and eluted in 70% acetonitrile with 0.05% acetic acid. Eluted peptides were vacuum centrifuge dried

and resuspended in 100 µl of 20 mM ammonium formate (pH 10) with 4% acetonitrile. Sample complexity was reduced by peptide fractionation using a Shimadzu HPLC system coupled to a photodiode array detector. Priming was done using a 1:1 ratio of mobile phase A (20 mM ammonium formate buffer, pH 10) to mobile phase B (80% acetonitrile, 20 mM ammonium formate) for 10 min at 1 ml/min before reducing to 5% mobile phase B. Peptides were loaded via a single partial loop injection, onto an Ascentis C18 HPLC column (octadecyl bonded phase; 150 mm x 4.6 mm, 5 µm; pore size 100 Å). Peptides were eluted at a flow rate of 1 ml/min using an initial isocratic low organic mobile phase (5% B) after which the organic phase was increased using a multi-step gradient up to 60% B over a total run of 75 min with 1 min fractions collected. Collected fractions were pooled to approximately similar peptide abundance by combining peptides of different retention times into 13 fractions and dried by vacuum centrifugation.

3.5.5. Mass spectrometry

Labelled samples were analysed (Central Analytical Facility, University of Stellenbosch) using a Dionex Ultimate 3000 RSLC nano LC (Thermo Scientific; Massachusetts, USA) system coupled to a Thermo Scientific Fusion Orbitrap Mass Spectrometer equipped with a Nanospray Flex ionization source. Peptides (1 - 2 µg) were loaded (mobile phase A: 2% acetonitrile with 0.1% formic acid) onto a C18 trapping-column (Thermo Scientific; 5 mm x 300 µm, 5 µm; pore size 100 Å) and a Luna C18 analytical column (Phenomenex; 350 mm x 75 µm, 3.6 µm). Samples were loaded onto the trap column at a loading-pump flow rate of 15 µl/min from a temperature controlled autosampler (7°C) for 5 min before eluting onto the analytical column. Peptide separation was performed at 50°C, at a flowrate of 500 nl/min using a non-linear gradient of 2 - 50% mobile phase B (100% acetonitrile with 0.1% formic acid) over 105 min. MS2 acquisition was performed using

monoisotopic precursor selection for ions with charge states between 2+ and 6+. Undetermined charge states and charge states > 24 were excluded. Dynamic exclusion was conducted with mass error tolerance of ± 10 ppm with isotopes excluded after 1 time. Intensity threshold was set at 5.0e4. Scans were collected in data dependent mode with a 3 second cycle time. Selected precursor ions were fragmented by higher-energy collisional dissociation set at a normalized collision energy (NCE) of 38% in the quadrupole mass analyser and then excluded from fragmentation once for 30 seconds. Fragment ions (MS2) were detected in the orbitrap mass analyser (resolution: 60 000), using centroid mode, with the automatic gain control (AGC) target of 1.0e5 and maximum ion injection of 120 ms.

3.5.6. Data processing and analysis

.Raw spectrum files were loaded into proteome discoverer version 1.4.1.14 (Thermo Scientific, USA) and spectra were filtered using a minimum and maximum precursor mass of 350 and 5000 Da respectively with a threshold peak count of 15. Precursor and fragment masses were set to 20 ppm and 0.02 Da respectively with a maximum of 2 missed tryptic cleavages permitted. Dynamic modifications allowed for oxidation of methionine as well as deamination of glutamine and asparagine. Static modifications included Carbamidomethyl +57.021 (C) Da, TMT 6-plex (K) Lysine +229.163 Da, and TMT 6-plex (N-terminus). Peak lists were searched against a UniProtKB/Swiss-Prot human database (*Homo sapiens*, Canonical sequences, November 2018, Sequences: 20 194) concatenated with the common Repository of Adventitious Proteins (cRAP) using a sequential alternating SequestHT/MSAmanda search engine schema. Amino acid modifications were iteratively added for each new spectral search cycle. Files from Proteome Discoverer software (.msf) were imported into Scaffold Proteome Software [266] for data validation using X!Tandem. Final spectrum and peptide matching

validation was conducted using Peptide Prophet and Protein Prophet Algorithms with the false discovery rates (FDR's) for protein and peptides set to 1% and 0.1% respectively. Relative quantitation was performed using the reporter ions quantifier built into Scaffold with the averaged D0 replicate samples set as the reference group. Proteins with missing values and not identified with at least 3 unique peptides were filtered out. Those passing validation were further analysed using InfernoRDN v. 1.1.7626.35996 (https://omics.pnl.gov/software/InfernoRDN; [232]), Perseus v. 1.6.7.0 software (Max Planck Institute of Biochemistry) [231] and RStudio [233]. For analysis conducted within Perseus and RStudio, data were transformed to log2(x) and analysed relative to the D0 experimental group. Global dataset statistics were conducted and group variations, standard deviations, protein ratios, and associated confidence scores were assigned for each protein group identified across TMT experiments. Hierarchical clustering was performed by clustering protein groups using Euclidean distances, and principal component analysis (PCA) was done using the Benjamini-Hochberg cut-off method with a 0.01% FDR. Volcano plots were generated using two-tailed t-tests and mean log2 fold change of 0.3 as well as an FDR cut-off of 0.05 used to assign significance (q-value). Correlation plots were generated using InfernoRDN, to allow all experimental groups to be relatively correlated, relative expression data from D0C, D14C, D21C and D28C were normalised to a reference group. Post normalisation, protein replicates were imported as relative-expression datasets and correlation plots generated. Annotations were assigned to protein accession numbers for gene ontology (GO), biological processes (BP), molecular functions (MF), cellular components (CC). Other analysis was conducted using GraphPad Prism windows software V7.0.0 (www.graphpad.com).

3.5.7. Spheroid fixation and immunofluorescent based staining

HepG2 cells grown in monolayers were harvested, as described above, and seeded (60000 cells/well) onto 12 mm coverslips coated with 500 μl Corning Matrigel Growth Factor Reduced (diluted 1/30 in pre-chilled DMEM) in 48-well plates. Cells at 70-80% confluence, were prefixed with 500 μl of 4% paraformaldehyde (PFA) for 10 min and then fixed in 4% PFA with 8% sucrose for 40 min. Coverslips were washed 3 times with PBS before cell permeabilization and blocking in B-PBT (1% Triton X-100, 10% FBS, and 4% bovine serum albumin in PBS) for 30 min. Samples were incubated with selected primary antibodies diluted in B-PBT (Albumin, HNF4α, and α-fetoprotein; Table 1) for 2 h followed by 3 washes in B-PBT before incubating with secondary antibody (Table 1) for 2 h. Whole spheroids were harvested by pipette aspiration and fixed in 4% PFA with 8% sucrose overnight. Spheroids were washed 3 times for 1 h in PBS on a plate shaker at 100 rpm. Permeabilizing and blocking was conducted as above in B-PBT for 2 h. Spheroids were incubated with selected primary antibodies (Albumin, HNF4α, and α-fetoprotein; Table 1) in B-PBT overnight followed by washing twice for 2 h in 0.2% PBT (0.2% Triton X-100 in PBS) and one wash in B-PBT for 2 h before incubating in secondary antibody (Table 1) in B-PBT overnight. Separate samples were LiveDead stained with fluorescein diacetate (FDA) and propidium iodide (PI) as described previously [176]. Positive controls for apoptosis induction were included as spheroids incubated in 10 μg/ml Puromycin for 16 h.

Table 3.2: Primary and secondary antibodies for immunofluorescent staining.

Primary antibody	**Species**	**Vendor**	**Dilution**
Albumin	Chicken	Abcam	1/100
HNF4α	Rabbit	Abcam	1/100
α-fetoprotein	Mouse	Abcam	1/50
Secondary antibody	**Species**	**Vendor**	**Dilution**
Alexa Fluor 488	Goat anti-chicken	Abcam	1/1000
Alexa Fluor 555	Donkey anti-rabbit	Abcam	1/1000
Alexa Fluor 647	Donkey anti-mouse	Abcam	1/1000

3.5.8. Tissue clearing and microscopy

Tissue clearing was performed using an adapted version of the ClearT2 method [267]. Briefly, antibody labelled spheroids were incubated, with agitation, in 25% formamide and 10% polyethylene glycol 6000 (PEG) in water. After 10 min the solution was changed to 50% formamide, 20% PEG and incubated for 2 h replenishing every 30 min until spheroid opacity diminished. Spheroids were washed briefly in PBS before mounting on a microscope slide, with a 0.12 mm spacer, using ProLong Diamond Antifade Mountant with DAPI overnight. Spheroids were then imaged using a Zeiss LSM 800 confocal microscope using a 20x objective. The aperture was set to 1 Airy unit and z-stack images were collected for each laser channel used (488 nm, 555 nm, 647 nm). Distances between z-stack image acquisitions were optimized per sample. Fluorescence and light microscopy were conducted using a Zeiss AxioVert A1 fitted with a Zeiss AxioCam digital camera. Images were analysed using Zeiss Zen Blue 3.0 software. Fluorescent signal from images was quantitated using the "Measure" tool built into the Zen Blue software package and normalised to relative fluorescence measurements to DAPI.

3.6. Contributions, funding, acknowledgments, conflicts of interest and data statements

Author Contributions: ADC, AAE and TH designed the study; Data were curated by AAE; Formal analysis was conducted by, AAE, MV; Funding acquisition, ADC and TH; Analysis and investigation, AAE and MV; Methodology design, AAE, MV, IvdB and TH; Project administration, AAE; The project was supervised by, IvdB, ADC and TH; Original manuscript was prepared by, AAE and TH; All authors contributed to the editing and revision of the manuscript.

Funding: The study was supported by the National Research Foundation (NRF) Thuthuka PhD Track Funding Scheme (Grant No. 87880), the NRF Innovation Masters and Doctoral Scholarship Program (Grant No. 107378), University of Pretoria Doctoral Bursary as well as the NRF National Equipment Program.

Acknowledgments: The authors would like to acknowledge Natasha Jeftha and express their utmost appreciation for her time, advice and support in securing funding for this study.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD024353 and 10.6019/PXD024353.

Chapter 4. Long-term drug cocktail exposure alters cellular phenotypes in HepG2 spheroid cultures

4.1. Chapter Contextualization:

The ability to recapitulate specific cellular states *in vitro* with acceptable accuracy is required when modelling disease or organotypic conditions. The dynamic temporal proteome changes observed in three-dimensional HepG2 cells were presented and discussed in Chapter 3 [182]. These data provide evidence for cellular adaptation in the context of transitioning from monolayer to spheroid culture environments, and the implications of the time course in culture on model applicability. Distinct differential expression was detected within the 2793 quantitated proteins, following spheroid culture for 28 days compared to the initiating monolayers was observed. Additionally, there were notable increases in expression of extracellular matrix, cell-junction/adhesion proteins as well as subsets of hepatic proteins indicative of a maturing hepatic phenotype. However, the data presented in Chapter 3 do not account for the influence of exogenous drugs within this system, which is essential if HepG2 spheroids are to serve as *in vitro* preclinical liver models. Typically *in vitro* assessments of hepatotoxicity or drug metabolism employ drug concentrations which far exceed their clinically relevant ranges [268]. While this may be necessary to achieve a measurable response, it limits the models' potential physiological inferences. Proteomics is an extremely useful technique for the sensitive detection and accurate quantitation of small changes in cellular protein expression [269], as has been demonstrated in results reported in Chapter 3. In this chapter the impact of extended exposure of HepG2 cell spheroids to a set of enzyme specific drugs in a low dose drug cocktail on the dynamic, time-dependent changes to the proteome was investigated. Cells were cultured in the presence of a drug cocktail containing seven drugs, which are metabolized by specific CYP450 enzymes *in vivo*, to assess whether the observed changes to the proteome had functional consequences on drug metabolism. This

would support the use of these cell models for hepatotoxic or metabolomic assessments, as postulated by others [227, 262, 270-272]. To assess possible proteomic and metabolomic changes while preserving translational relevance, the seven drugs in the cocktail were used at their typical clinical maximum observed concentration (C_{max}), that had been shown to be well below cytotoxic thresholds as established in the earlier experiments conducted and reported in Chapter 2.

4.2. Materials and Methods

4.2.1. Cell culture and long-term drug exposure of spheroid cultures

Cultures of HepG2 spheroids were seeded, grown, and maintained using the hanging drop method as described in Chapter 2. For the assessment of proteomic changes occurring due to long term drug exposure in spheroid cultures, drug cocktail culture media was made by dilution of 50 mM drug stock solutions prepared previously to each drug's respective Cmax as detailed in Table 4.1 below. Drug dilutions were made directly into cell culture media supplemented with 10% FCS and 2 mM GlutaMax to match conditions of spheroids cultured in earlier chapters. Cells harvested from monolayers for seeing spheroids (20 000) were resuspended directly in either normal complete media or drug cocktail supplemented media. Spheroid cultures, as control or drug treated groups, were cultured in parallel over the same time course of 28 days with samples collected for proteomic analyses at days 14, 21 and 28 in culture. For ease of reference, spheroid cultures exposed to drug cocktail are referred to as the "induced" denoted by the "I" suffix in downstream abbreviations. Medium exchanges were conducted as described in Chapter 2, with the induced groups receiving fresh drug cocktail supplemented complete DMEM media at each exchange. The DMSO concentration in the media for the induced groups was lower than ~ 0.03% at which concentration no cytotoxic effects were observed.

Table 4.1: Drugs and respective clinical Cmax concentrations which were used for culture

Drug Name	Approximate C_{max} (ng/mL)	Clinical Dose (mg)	References
Artemisinin	60.0 ± 32.5	20 x 4 (oral)	[273, 274]
Buspirone	1.13 ± 0.15	10 (oral)	[275, 276]
Dextromethorphan	2.9	30 (oral)	[277, 278]
Diclofenac	524 ± 222	100 (oral)	[279]
Midazolam	71 ± 25	5 (nasal)	[280, 281]
Omeprazole	283 ± 113	20 (oral)	[282]
Phenacetin	2100 ± 1300	1000 (oral)	[283]

4.2.2. Quantitative mass spectrometric proteomics

All the materials and methods for processing of the drug exposed HepG2 spheroid samples for mass spectrometric analysis, including sample collection, spheroid lysis, protein quantitation, reduction and alkylation, trypsin digestion, TMT-labelling, desalting by SPE clean-up, high-pH reverse-phase HPLC fractionation and LC-MS/MS analyses were conducted as described in Chapter 3 where the culture time course proteome changes were assessed.

Isobaric peptide tagging of experimental groups were differentially labelled as shown in Table 4.2. Samples were labelled according to the following convention as DXXER where D represents days in culture with XX an integer placeholder for the day in the 28-day culture time course, E is the drug exposed experimental condition with C indicating non-exposed controls or I indicating drug cocktail exposed experimental groups, and R representing the respective biological replicate number. Samples indicated as D0 were the reference monolayer cultured controls used to initiate the spheroid cultures. Samples denoted as "Pool" comprised of equivalent amounts of protein (15 μg) from each sample within respective replicate tag groups included in the pooled set. Label assignments for TMT isobaric tags were subject to forward,

reverse and partial randomisation labelling to account for potential tag-sample affinities or reaction biases.

Table 4.2:Tandem mass tag sample labels for peptides derived from various experimental conditions

Tag number	Replicate 1	Replicate 2	Replicate 3	Replicate 4
126	D14C1	D14C2	Pool	Pool
127	D14I1	D14I2	D0C2	D0C3
128	D21C1	D28C3	D28I2	D21I2
129	D21I1	D28I3	D28C2	D21C2
130	D28C1	D0C1	D21I3	D14I3
131	D28I1	Pool	D21C3	D14C3

4.2.3. *Proteomics data processing and analyses software/packages*

Data processing was conducted as described in Chapter 3 with respect to search engines, data filtering and software pipelines. However, the final bioinformatics analysis pipelines were modified, and are discussed in more detail below. Briefly, relative quantitation was conducted using the reporter ion quantifier built into Scaffold V4.10 [266] for each confidently identified protein. Protein lists with relative quantitation, normalized to D0 values in Scaffold, were imported into R and R Studio for further processing [284, 285] normalised log(2x) fold change ratios were scaled up from continuous to discrete values compatible with downstream analyses, by applying a 10^4 scale factor to the entire dataset. The detection of differentially expressed proteins (DEPs) was conducted using the DESeq2 package [286] and the Perseus computational platform for proteomics [231]. Biological inferences were made using a combination of the Gene Ontology (GO) databases [287], Molecular Signature Database (MSigDB) [288], The Reactome Knowledgebase [289] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [290]. Other miscellaneous statistical and bioinformatics packages imported into R included, Tidyverse [291], DplyR [292],

ggplot2 [293], FGSEA [294], GSVA [295], Pathview [296], Gage [297], Goplot [298], pheatmap [299] and EnhancedVolcano [300].

4.2.4. Immunohistochemistry and confocal microscopy

Spheroid fixation, immunohistochemistry processing and confocal microscopy were conducted as described in Chapters 2 and 3. However, for tissue clearing the ClearT2 method was replaced with an ultrafast optical clearing method (FOCM) originally described by Zhu *et al.* [301] which makes use of inexpensive and readily available, nontoxic reagents. The FOCM method has been shown to optically clear tissue slices of up to 300 μm in thickness, in as little as 2 min, without compromising the tissue morphology or causing immunofluorescent quenching. Briefly, FOCM reagent was prepared by dissolving urea to 30%, D-sorbitol to 20%, and glycerol to 5% in terms of w/v in DMSO, all being highest purity reagents purchased from Sigma Aldrich (St. Louis, MO, USA). Optical clearing by FOCM was performed by transferring immunofluorescent labelled spheroids into FOCM reagent and gently mixing by aspiration to ensure complete exposure to the clearing solution. Spheroids were then incubated for approximately 5 min becoming mostly transparent. Antibodies used for the detection of hepatic marker proteins are summarised previously in Chapters 2 and 3.

4.2.5. Assessment of cytochrome P450 drug metabolism activity in HepG2 spheroid cultures

Metabolism of commercial drugs is largely facilitated by hepatic CYP 450 enzymes. These enzymes usually facilitate the phase 1 metabolism of drugs with specific chemical moieties although CYPs generally show overlap in chemical classes of drugs they can metabolise. Using this relationship, it is possible, under controlled conditions to attribute the formation of a specific metabolite to the activity of a specific CYP proteoform. LC-MS/MS based analysis combines high sensitivity, selectivity, and

quantitative accuracy with the capacity for simultaneous quantification of multiple physiochemically distinct compounds. This makes LC-MS/MS an indispensable technique for assessing the biotransformation capacity of candidate *in vitro* hepatic models. A metabolomic study, using a selection of probe drugs from the Geneva phenotyping cocktail [177] was conducted to assess the temporal impact of three-dimensional culture on the metabolic competence of HepG2 cells and to establish the functional consequences of extended cell culturing in the presence of a drug cocktail. The list of probe drugs used in the metabolic test drug cocktail and their respective CYP450 targets are listed in Figure 4.1 and Table 4.2 below.

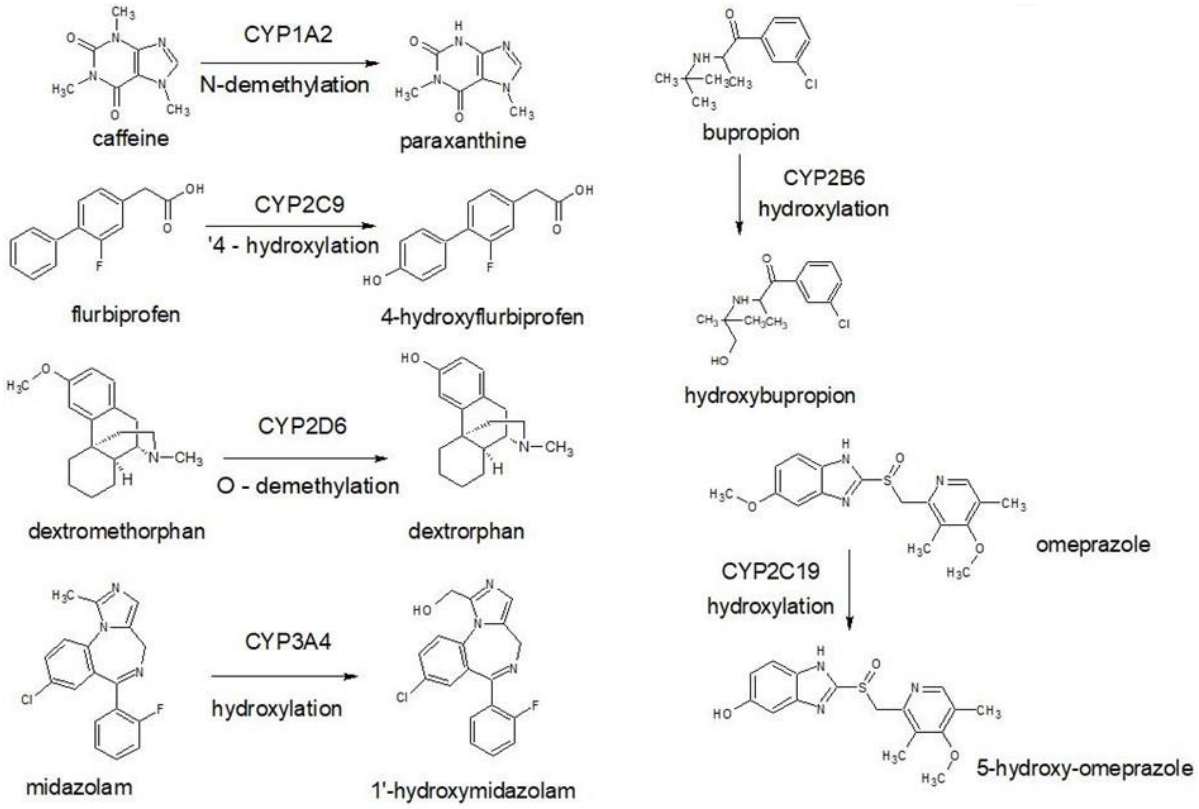


Figure 4.1: Metabolic pathways of the 7-drug Geneva phenotyping cocktail. Adapted with permission from [302]

Table 4.3: Cytochrome P450 probe drugs, their respective enzyme targets, and metabolites

Probe drug	Target CYP enzyme	Metabolite
Caffeine (CAF)	CYP1A2	Paraxanthine
Bupropion (BUP)	CYP 2B6	4-Hydroxybupropion
Flurbiprofen (FLB)	CYP 2C89	4-Hydroxyflurbiprofen
Dextromethorphan (DEX)	CYP 2D6	Dextrorphan
Omeprazole (OPZ)	CYP 2C19	5-Hydroxyomeprazole
Midazolam (MDZ)	CYP 3A4	1-Hydroxymidazolam

To increase spheroid numbers for these metabolic investigations, the hanging drop culture method was replaced by the 3D micro-mould culture method (described in Chapter 2). The 3D micro-mould allows for the simultaneous generation of 81 spheroids/well and required less frequent growth medium exchanges. The use of this culture method is preferred when requiring high throughput culture, to maximise spheroid generation capacity for metabolomic studies. The reduced requirement for medium exchanges and lack of a hanging-drop is also desirable for improved handling and culture maintenance. Equivalent HepG2 spheroid cultures were cultured over the same 28-day time course as control or drug cocktail exposed groups, using the drug cocktail composition and drug concentrations shown in Section 4.2.1, Table 4.1 above. On selected terminal days (Days 14, 21 and 28) spheroids (n=81) were collected and transferred to sterile Eppendorf tubes, washed 3 times with 1 mL PBS to remove media and any residual drug cocktail for the induced group. Spheroids from both control and induced groups were transferred to new Eppendorf tubes and exposed to 1 mL of media containing drug probes from the Geneva phenotyping cocktail. All the probe drug were added at a fixed concentration of 5 μM to avoid loss of specificity for their target CYP and non-specific metabolism [303]. Separate sets of spheroids were incubated at 37°C for 3, 6 and 24 h for 3 biological replicates. After the

incubation period, spheroids were pelleted by briefly pulse centrifuging and the media supernatant collected and transferred to a new tube.

4.2.6. Analyte extraction from biological matrices in preparation for MS analysis

Samples prepared and stored as described in Section 4.2.5 above were thawed at 4°C then desalted and concentrated using an optimized solid phase extraction (SPE) method. Briefly, samples were thawed, and diluted with 10 volumes of MS-grade water. Cartridges (Varian Bond Elute Plexa, 60 mg bed, 1 cc vacuum type cartridges) were activated and conditioned with 1 mL of 100% methanol, equilibrated twice with 1 mL of MS-grade water and slowly loaded (10 x 1 mL) onto the cartridge without allowing the media to dry. Loaded analytes were then washed with 3 mL of MS-grade water before slowly eluting with 2 mL of 80:20 methanol:ACN and collecting into clean Eppendorf tubes. Eluates were vacuum dried in a CentriVap Benchtop Vacuum Concentrator (Labcono, Kansas City, MO, USA), and stored dry at -20°C until analysis. Prior to analysis, died samples were resuspended in exactly 100 µL mobile phase (5% mobile phase B in mobile phase A) spiked with 2.5 µM of Acetaminophen (ACAP) serving as the internal standard. Samples were solubilized under these conditions to match the starting conditions of the LC-MS/MS method.

4.2.7. Metabolic profiling by mass spectrometric assay

A method for the detection and quantitation of CYP probe drugs and their metabolites, was adapted from a validated in-house developed method [302]. The LC-MS/MS system consisted of an Agilent LC system (Agilent Technologies, Palo Alto, CA, USA), equipped with a 1290 Infinity Binary Pump, Infinity II autosampler, multicolumn thermostat and solvent degasser. This LC system was coupled to a Sciex 4000QTrap triple quadrupole mass spectrometer, equipped with a Turbo-V® electrospray ionization (ESI) source (Sciex, Concord, Canada). Analyst™ Software, version 1.7.2 (Sciex, Concord, Canada), was used to operate the system, manage the

method optimization, data acquisition, and perform quantitative data analysis. Analyst device driver was used for the method integration and management of the Agilent LC method. Quantitation of probe drugs and their metabolites was achieved using a targeted multiple reaction monitoring (MRM) method. Instrument tuning for each analyte was conducted via direct analyte infusion of 10 μg/mL solutions using a Harvard syringe pump (Harvard Apparatus, Hollinston, MA, USA) with a flow rate of 10 μL/min. Precursor ion scans (Q1) were conducted for all probe analytes in both positive and negative mode, with the initial acquisition settings as follows: ion spray voltage (ISV) at 5500 V or -4500 V, curtain gas (CUR) set to 23 psi, turbo heater temperature at 200°C with the nebuliser gas flow rate kept at 36 psi and the heater gas flow rate at 35 psi. The delustering potential (DP) was optimised for each analyte in both positive and negative modes respectively.

Ions from the Q1 scans with the highest intensity for molecular masses (M) of [M+1] in positive mode or [M-1] in negative mode were selected for optimization of transition pairs in product ion scans (MS2). The collision energy (CE) in Q2 was ramped to determine the optimum parameters for fragmentation of precursor ions to product ions (detected in Q3). Peripheral method parameters (source gas parameters) were optimized in tandem for each transition pair. A multi-step gradient separation of analytes was achieved using a Kinetex™ Biphenyl column (100 × 2.1 mm, 2.6 μm particle size) with mobile phases consisting of 10 mM ammonium formate for mobile phase A and an organic mobile phase B consisting of ACN:Methanol in a ratio of 20: 80 with 10 mM ammonium formate. The method details for the HPLC and MS acquisition are detailed in Tables 4.4 and 4.5 below, respectively. Static parameters following initial optimisation for positive mode were as follows: Curtain gas 23 psi, collision gas set to medium, ISV to +5500 V, source temperature was maintained at 450°C, ion source gases were set to 35 and 36 psi for G1 and G2 respectively, interface heater on and entrance potential and collision cell exit potential both set to 10 V. Parameters for negative mode were kept identical for all gasses and temperature, ISV

was set to -4500 V and entrance potential and collision cell exit potential both set to -10 V. Five microlitres of resuspended samples containing metabolites were injected and eluted from the column using an isocratic flow rate of 250 μL/min and gradient of increasing mobile phase B as per the parameters detailed in Table 4.4 below. Data were collected using a targeted MRM method for presence of the parent compound and paired metabolite transition pairs noted in Table 4.5 below. Area under the curve (AUC) was measured and used for relative quantitation of parent-metabolite pairs with the increases or decreases in metabolite formation inferred relatively as normalised to parent compound

Table 4.4 Details of chromatographic gradient used in HPLC method

Time [min]	A [%]	B [%]	Max pressure limit [bar]
0.00	90	10	800
1.00	90	10	800
1.25	68	32	800
13.00	10	90	800
14.50	10	90	800
15.00	90	10	800
17.20	90	10	800

Table 4.5:Optimised MS fragmentation parameters for analytes and internal standard

Analyte	***Probing***	***Monoisotopic mass***	***Q1 Mass (Da)***	***Q3 Mass (Da)***	***Dwell time (msec)***	***ID***	***DP***	***CE***	***Mode***
Caffeine	CYP1A2	194	195.3	138.2	100	CAF	20	25	Positive
Paraxanthine		180.1	181.1	124.2	100	PAR	70	27	Positive
Bupropion	CYP2B6	275.1	240.4	131.3	100	BUP	20	50	Positive
Hydroxybupropion		255.1	256.4	238.1	100	OHBUP	50	52	Positive
Omeprazole	CYP2C19	345.1	346.3	198.1	100	OPZ	25	30	Positive
Hydroxyomeprazole		361.1	362.1	214.4	100	OHOPZ	50	15	Positive
Dextromethorphan (1)	CYP2D6	271.1	272.4	147.4	100	DEX	90	50	Positive
Dextromethorphan (2)	CYP2D6	271.1	272.4	171.5	100	DEX2	90	50	Positive
Dextrorphan		257.4	258.4	157.2	100	DTP	80	45	Positive
Midazolam	CYP3A4	441.1	326.3	291.4	100	MDZ	80	35	Positive
Hydroxymidazolam		341.1	342.2	324.1	100	OHMDZ	89	29	Positive
Flurbiprofen	CYP2C9	244.1	242.9	198.7	100	Flu	-50	-50	Negative
Hydroxyflurbiprofen		260.1	259	215.1	100	4OH-Flu	-50	-20	Negative
Acetaminophen (IS)	NA	151.1	152.17	110.1	100	APAP	50	30	Positive

4.3. Results and Discussion

Many candidate *in vitro* hepatic models assess hepatic marker gene expression at the level of mRNA transcript or the secretion of classic hepatic marker proteins which can be easily misconstrued to represent physiological functionality [227, 262, 272, 304]. While these parameters are important to characterize these models, inferences that the models robustly replicate *in vivo* functionality should be made with care. Currently, no *in vitro* model is able to accurately replicate the highly complex interactions involving multiple biological systems seen *in vivo*. Global hepatic function is complex and relies on proteome crosstalk between multiple cell types present in the liver and the plasma. Even when constraining the model to represent one specific cell type, creating inferences from a limited panel of typical marker proteins would be

meaningless without contextualizing these markers against the global biological background for that cell type. In earlier chapters it was shown that HepG2 spheroids progressively change the expression of many protein groups compared to monolayer counterparts by simply transitioning from two-dimensional to three-dimensional cultures. This change in the proteome occurs while maintaining and enhancing, relative to monolayer, the expression of several typical hepatic marker proteins (Chapter 3, Figure 3.4) while progressively depositing extracellular matrix proteins (Chapter 2, Figure 2.12 and Chapter 3, Figure 3.5). Here the analyses have been expanded to include and characterise the proteomic and selected metabolomic changes in spheroids cultured for an extended time (28 days) in the presence of an induction drug cocktail relative to equivalent non-induced spheroids.

4.3.1. Comparison of the proteomic changes in long-term induced vs control HepG2 spheroids

Successful isobaric labelling and quantitative proteomic analyses successfully identified over 5000 proteins, filtering data as conducted previously reduced the dataset to 4817 proteins which were present in all replicates with at least 3 unique peptides. Unsupervised hierarchical clustering analysis based on changes in protein abundance is useful for simultaneously identifying the similarities and differences between the global proteome of experimental groups. Replicates of HepG2 spheroids clustered distinctly according to both culture timepoint and to drug exposure conditions. In general, spheroid experimental groups trended similarly in the progressive increased or decreased expression of specific pathway related proteins within their proteomes compared to the monolayers used to initiate the spheroids. The proteome of these monolayers served as the mean protein group level normalization point for all changes observed. Pronounced proteome changes were evident within the control spheroid group, which appeared to be dependent on the elapsed time in culture. Clustering of induced groups, while still evidently different

compared to monolayers, showed fewer time dependent changes from Day 14 through to Day 28 than for the control spheroids.

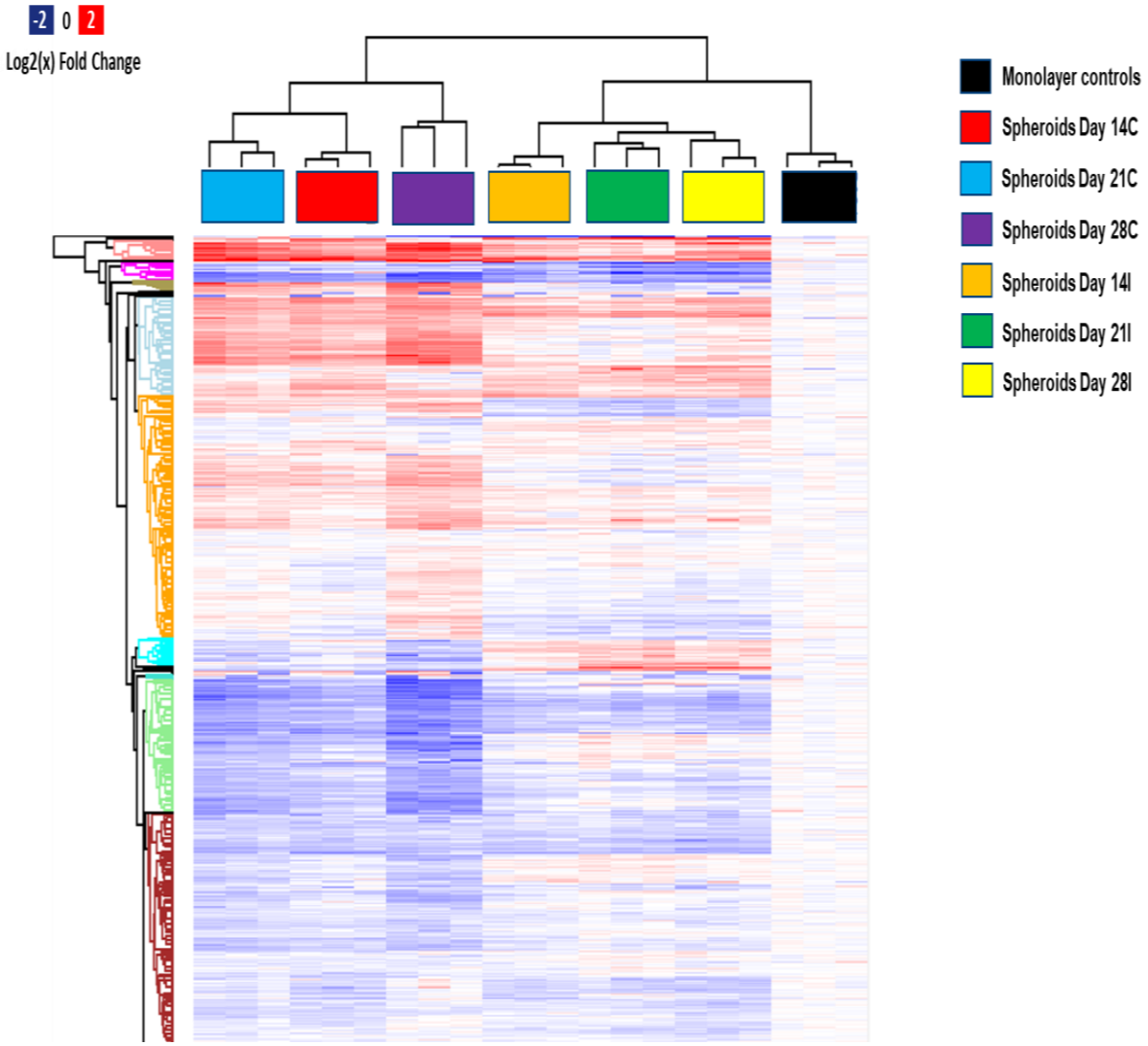


Figure 4.2:Hierarchical clustering of proteomic data from spheroid and monolayer experimental groups

The progressive changes in the proteome are also clearly visualized when examining the relative contributions to cohort variance as illustrated in the principal component analysis (PCA) plots (Figure 4.2). While all spheroid groups resolve from monolayer groups, clustering also resolved according to control or drug exposure conditions. A

progressive culture-time related difference could be seen (apparent up- and down-regulation with time) within the control spheroid groups.

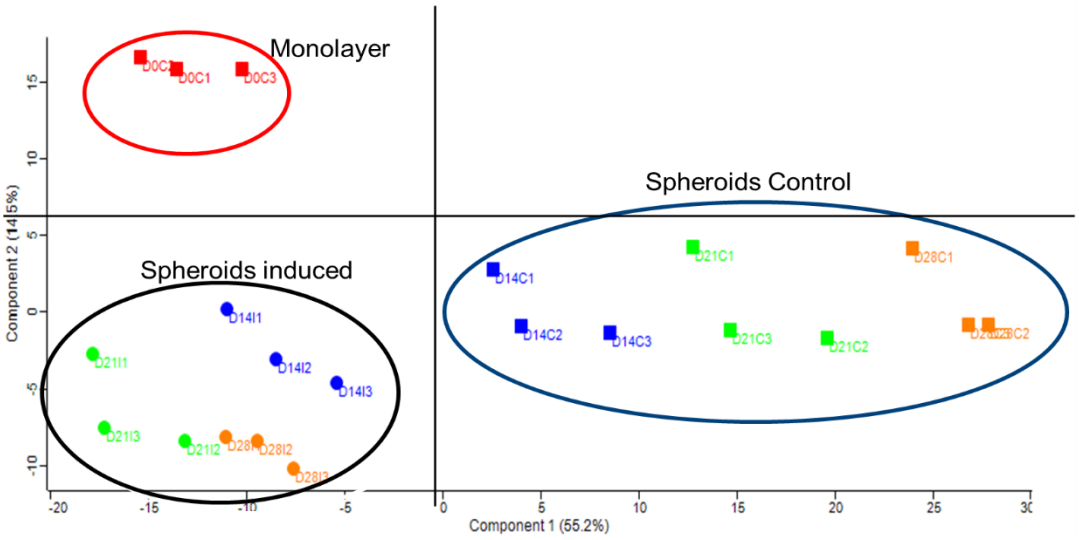


Figure 4.3: A PCA loading plot of HepG2 cell replicate proteomes for monolayers (red ring), drug induced spheroids (black ring), and control groups (blue ring), sample timepoints are denoted by colour and shape, for D0 monolayers (red squares), D14C and D14I spheroids (blue squares and circles respectively), D21C and D21I spheroids (green squares and circles respectively) and D28C and D28I spheroids (orange squares and circles respectively).

Another informative method for visualizing global trends across large datasets is though the generation of correlation matrices, whereby the agreement of change between datapoints of common identities across experimental replicates are plotted against each other. Pearson correlations (Figure 4.3) demonstrated the highest correlation (approaching 1) between biological replicates of the same time points and within the same experimental grouping (control or induced). Interestingly, while the correlation between control and induced spheroids was reasonably high at the beginning of the incubation time course, with correlation coefficients of ~0.6 - 0.75, these became progressively less correlated as incubation time increased. It could be postulated that while control and induced spheroid groups differ within each respective proteome relative to monolayers, the protein cohorts which are

differentially expressed across experimental grouping could reflect distinct functional differences.

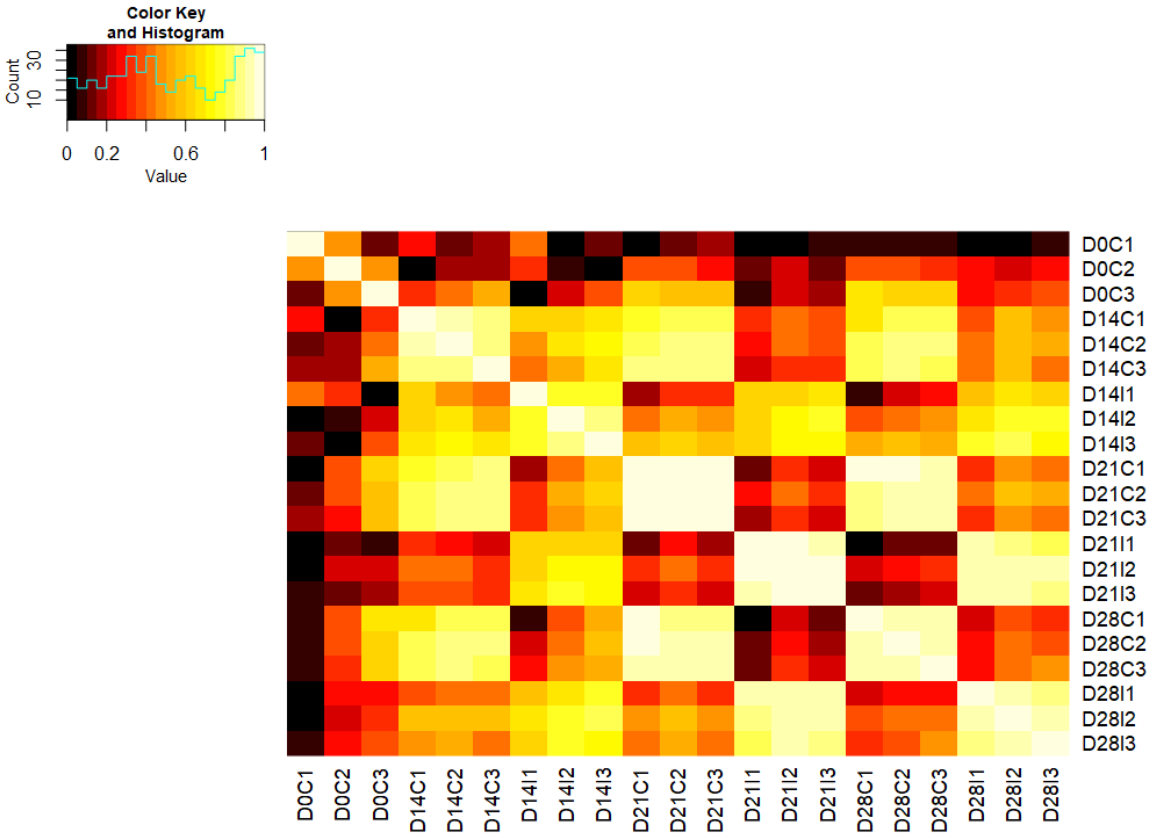


Figure 4.4: Correlation plot of Pearson correlations between experimental group

Understanding the reproducibility and proteomic concordance between biological replicates of spheroids across different time courses is important to validate the model's robustness and suitability of both the biological and technical assay approaches utilized. Understanding the probable mechanisms that drive these proteomic changes, as well as the effect of these changes have on cellular function are relevant objectives of this study. To this end, the functional and biological nature of the changes observed were characterized using the inferred protein identity and the relative changes in specific protein abundance that could be assigned to known biological pathways.

4.3.2. Bioinformatics analyses for high level biological inferences from proteomic data

The progressive development of quantitative methods has seen proteomics become a powerful tool for the analysis of biological systems. These methods can be leveraged to understand different aspects of systems biology such as disease progression or prognosis, biomarker identification as well as elucidating the proteomic outcomes of genomic and transcriptomic perturbations [305]. Classically these quantitative experiments make use of a hierarchical data structure, that is, intensity information from summed or median peptide spectrum matches are inferred into amino acid sequence of a peptide then the cumulative or median peptide abundances used to estimate matched protein group abundances. The subsequent analysis steps are critical for gaining insight into the biological underpinnings which determine shift in cell functions or phenotypic states. This process usually begins with statistical methods for determining the differentially expressed proteins (DEPs) across experimental groups. Despite many technological advances in the field of proteomics, lack of standardized methods for data analysis leaves bioinformatic approaches open for end user discretion. Proteomics datasets come with several unique challenges, including a potentially high number of missing values, batch and instrument bias effects, as well as the high degree of potential technical but especially biological variances. As such, computational tools and software packages have been developed for data normalization and imputation of missing values to meet statistical processing requirements in quantitative proteomics workflows. Smaller datasets in particular pose a challenge with missing values as the reliability of imputation becomes questionable, yet this phenomenon is less pronounced as datasets become larger [306]. The students t-test and analysis of variance are commonly used in proteomics for detecting DEPs even though both these tests carry well known limitations, which again, are more pronounced in smaller datasets [307].

In parallel to the technological strides made in proteomics fields, similar progress has been made in measuring gene expression. RNA sequencing (RNA-seq) has become the predominant tool for measuring gene expression, boasting high degrees of coverage, depth, throughput, and reproducibility. To meet the increasing demands of RNA-seq data analyses, several statistical pipelines have been developed, with some of the more widely adopted being SAM, DESeq2, Limma and EdgeR. These pipelines have seen notable success with some already transferred directly to proteomics workflows [307]. Some common issues with the widespread adoption of these methods for use in proteomics data analysis are the inability to handle missing data and the production of continuous data instead of discrete integer values as in the case of quantifying transcript reads. To circumvent the issue of missing values, a variety of data imputation strategies have been developed, such as the k-nearest neighbour (KNN), maximum likelihood estimate (MLE), and single value decomposition (SVD) being among the more widely used [308]. Generally, the nature of the missing values relative to the dataset informs the most appropriate type of imputation strategy [309], however, paired differences of quantification strategies and imputation methods may produce variable identification of DEPs and support false conclusions in the analysis of proteomics data.

During the initial filtering phase of proteomics data analysis pipelines, the most appropriate method of dealing with missing data needs to be decided. In the analysis used in this study, if identified proteins were not present in all replicates with at least 3 unique peptides and 100% confidence they were removed from the data set. While this substantially reduced the size of the dataset it improved the confidence in downstream analysis and mitigated the possible variance contribution arising from imputation. Additionally, by removing these values, the limitations of software package to deal with missing values were also avoided. Previously the Perseus proteomics software package for the identification of DEPs and further bioinformatics analysis was used. In Perseus, the assignment of DEPs is based on the t-test method.

DESeq2 a statistical package for the detection of differentially expressed genes (DEGs) in RNA-seq data has recently been shown to be suitable for the analysis of label free proteomics data [309-311]. The use of DESeq2 in isobaric tagged proteomics has been limited due to the software requirement for discrete integer values. To circumvent this a global scaling method was applied, which served a dual the purpose of converting the continuous relative abundance data into integer values and to rescale the dataset to within the normal RNA-seq data range. An additional benefit of using DESeq2 for analysis is that there are many downstream statistical packages built into R that accept the DESeq2 object as an input which conveniently expands the available analyses methods without the need for porting data across multiple analyses platforms.

Assessment of DEPs across spheroid experimental conditions and selected time course points was informative for understanding not only the degree of proteomic change conferred by each experimental culture condition, but also for understanding the holistic effect of these changes. When assessing DEPs, there are several considerations which must be carefully made. Firstly, statistical thresholding which generally serves to ensure that an appropriate statistical method is utilized. DESeq2 uses a population size factor normalized negative binomial generalized linear model to assign differential expression [286], supported with a Benjamini Hochberg FDR correction [312, 313]. The second consideration is that of biological thresholding, which relates to the fold change cut-offs, and are often subjective and disputable. While it is common that widely differential gene expression in transcriptomic studies can result in differences between experimental conditions with fold changes of orders of magnitude apart, this is typically not the case in proteomics data, with the exception being seen in knock-in overexpression systems or knock-out expression inhibition systems.

Comparisons of DEPs identified in day 14 spheroid groups compared to monolayer

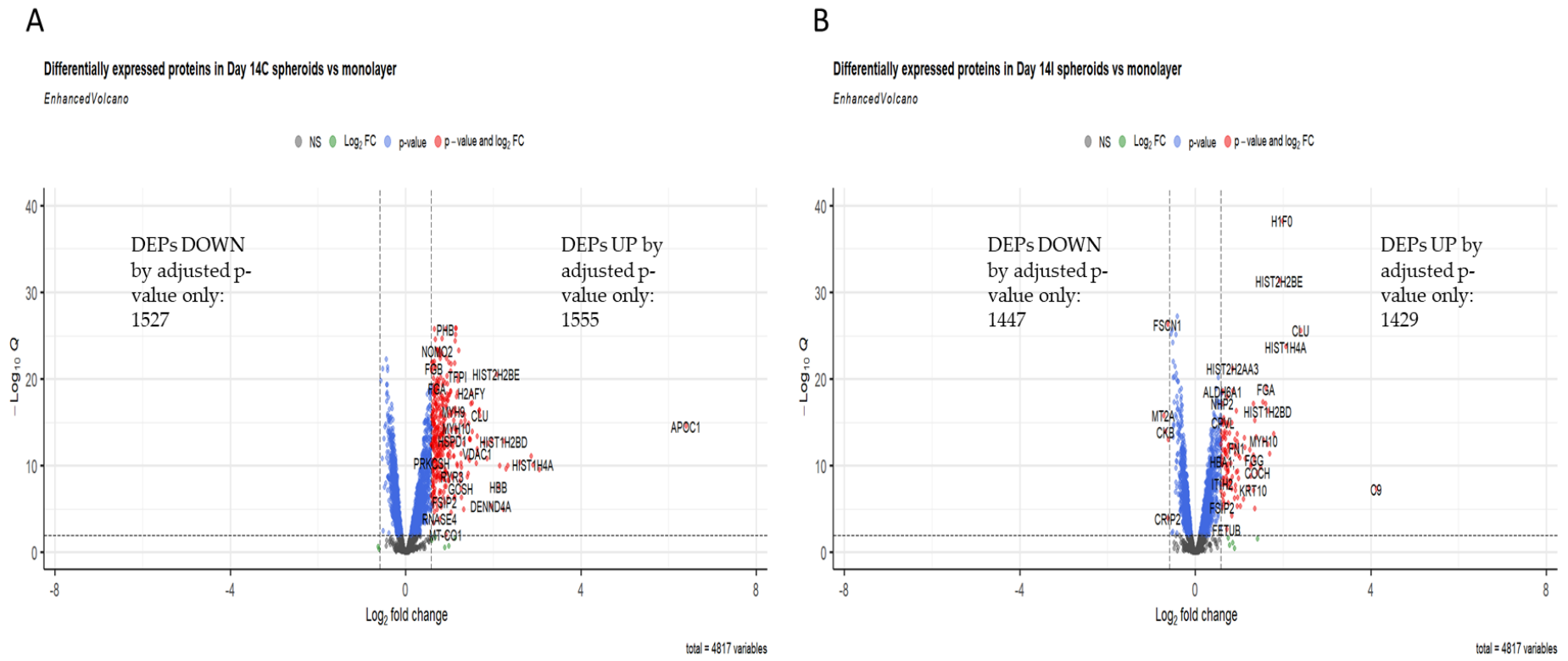


Figure 4.5:Volcano plots contrasting DEPs between A) D14C spheroids vs monolayer and B) D14I spheroids vs monolayer.

Comparisons of DEPs identified in day 21 spheroid groups compared to monolayer

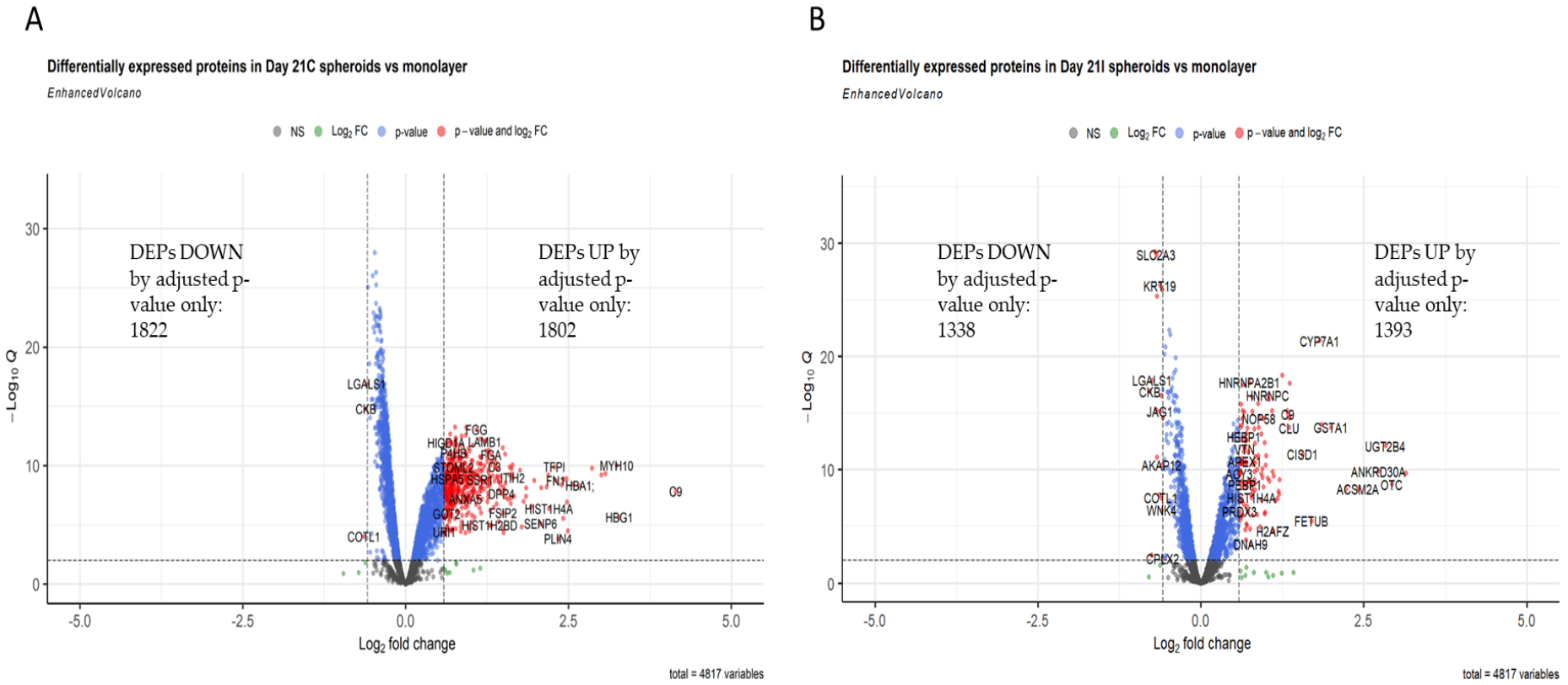
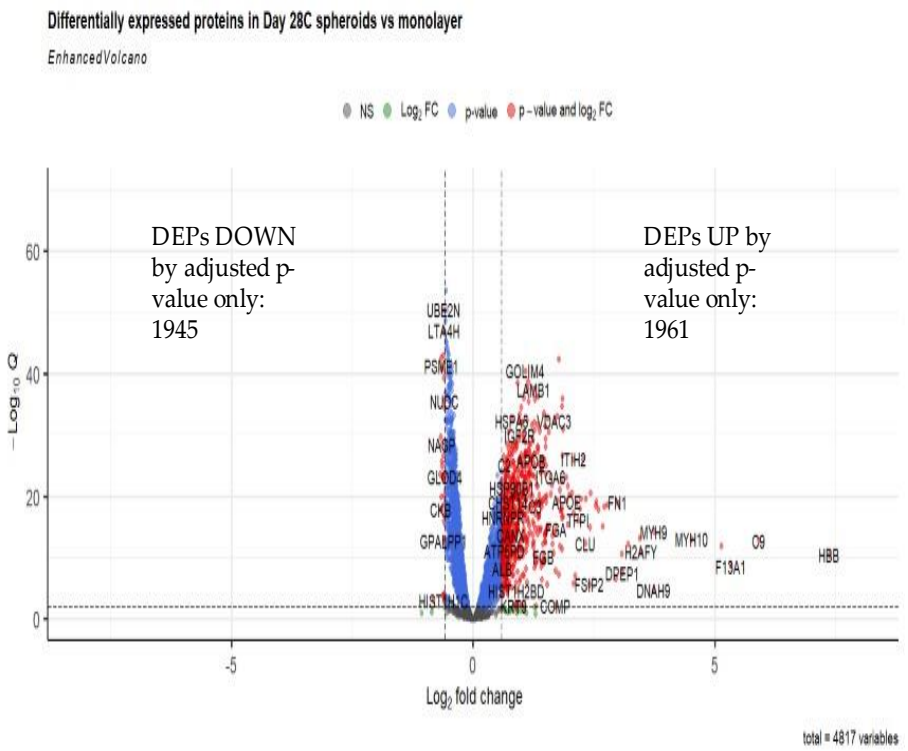


Figure 4.6: Volcano plots contrasting DEPs between A) D21C spheroids vs monolayer and B) D21I spheroids vs monolayer

Comparisons of DEPs identified in day 28 spheroid groups compared to monolayer

A



B

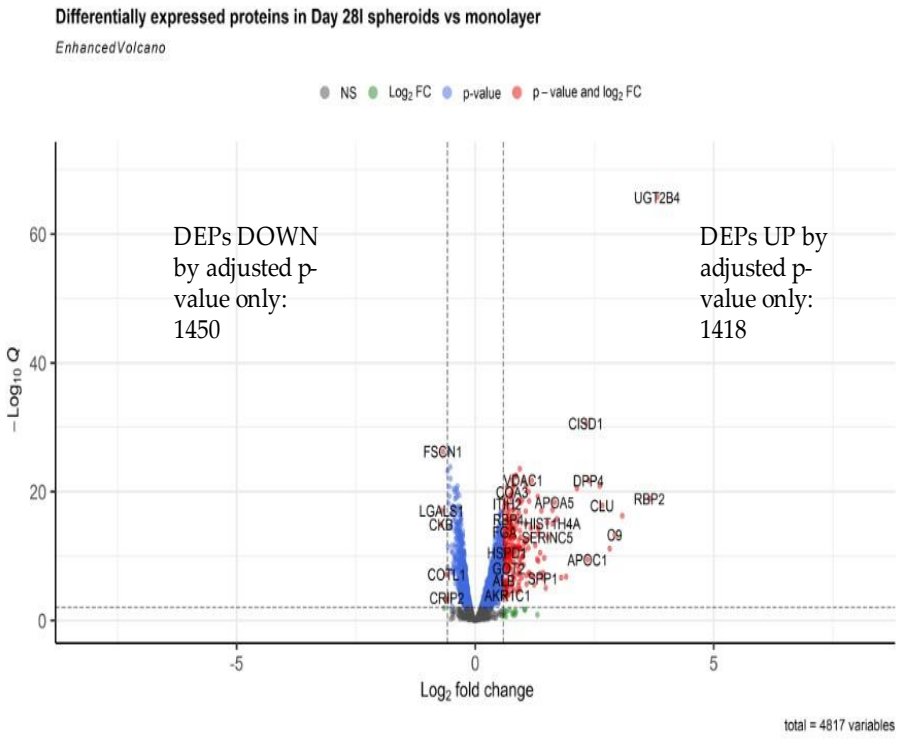


Figure 4.7:Volcano plots contrasting DEPs between A) D28C spheroids vs monolayer and B) D28I spheroids vs monolayer

Figures 4.4-4.7 show volcano plots of DEPs between same time point control and induced spheroids showing the expression changes relative to the initial monolayer counterparts at the three sampling timepoints over the 28 day culture time course, generated using the EnhancedVolcano R package [300]. Here, the monolayers were set as the reference group within the DESeq2 object using the factor relevel function. Therefore, changes in the spheroid groups were expressed as their relative fold changes compared to the initial monolayer cultures used to initiate the spheroid cultures, i.e. positive fold changes (positive values in red to the right) are due to higher expression in spheroids, and negative fold changes (in red to the left) are due to lower expression in spheroids. In general, all the spheroid experimental groups showed similar positive and negative differential expression compared to monolayers when considering only the statistical threshold (p-value). However, when also considering the relative fold changes (p-value and Log_2 FC), a larger number of proteins were significantly differentially expressed in the untreated control spheroid groups seen in Figure 4.4 A, C and E. When superficially examining the types of proteins increasing between control and drug induced spheroid groups, there is overlap in upregulated protein ID's. Notably, in the earlier time points, multiple histone proteins were increased in both treated and untreated spheroids. Yet, disparities became evident across control and induced spheroid groups as the number of DEPs increased progressively over the time course in only control spheroid groups. Proteins showing increased differential expression in the induced groups included several with metabolic activity such as alcohol dehydrogenases (ALDH), ornithine transcarbamylase (OTC), Cytochrome P450 7A1, and UGT2B4 among others, while in the control groups various structural proteins showed increased differential expression, such as several non-muscle myosins (MYH9&10), laminin (LAMB1) and fibronectin (FN1) as well as some hallmark hepatic protein markers such as albumin (ALB) and apolipoprotein B (APOB).

4.3.3. Investigation of differentially expressed proteins across control and drug treated spheroid cultures

The data presented above and in Chapter 3 highlight two variables which contribute to proteomic change in HepG2 spheroid cultures. The first being the temporal element where changes are proportional to time spent in culture, though this is continuous in control spheroids with evidence suggesting stasis between day 14-28 in induced groups. The second variable being long-term exposure to the drug cocktail. Given the observable disparity in DEPs after exposure to relatively low levels of drug cocktail, differences between DEPs from control and induced spheroids at identical culture time points were characterized. To achieve this, the DplyR package in R was used to subset the dataset to enable direct comparisons between groups of interest. Volcano plots were used and unsupervised hierarchical clustering done using the top 50 proteins (Figure 4.5) with highest Z-score variances using the pheatmap R package [299]. Contrasts were generated using the relevel function, as done previously, with the control spheroids group set as the reference level factor i.e. positive values are indicative of higher expression in induced spheroids and negative values are indicative of higher expression in control groups. Consistent with what was observed previously, when comparing differences using significance thresholds alone (p-value), similar numbers of DEPs are noted between the experimental groups, but when the fold change contributions (p-value and Log_2 FC) are included, the control spheroids show a greater magnitude of change. Unsupervised clustering of control versus induced groups indicated similar trends, (Figures 4.8-4.10) with variance contributed by upregulation being largely attributed to the control group spheroids. This is also noted when comparing these groups relative to baseline monolayer (D0) proteomes. However, as demonstrated by the paired volcano plots in the same figures drug treated spheroid groups also show progressive changes of different protein groups over time, though with a comparatively lower overall fold change.

Comparisons of DEPs identified in 14C vs D14I spheroid groups

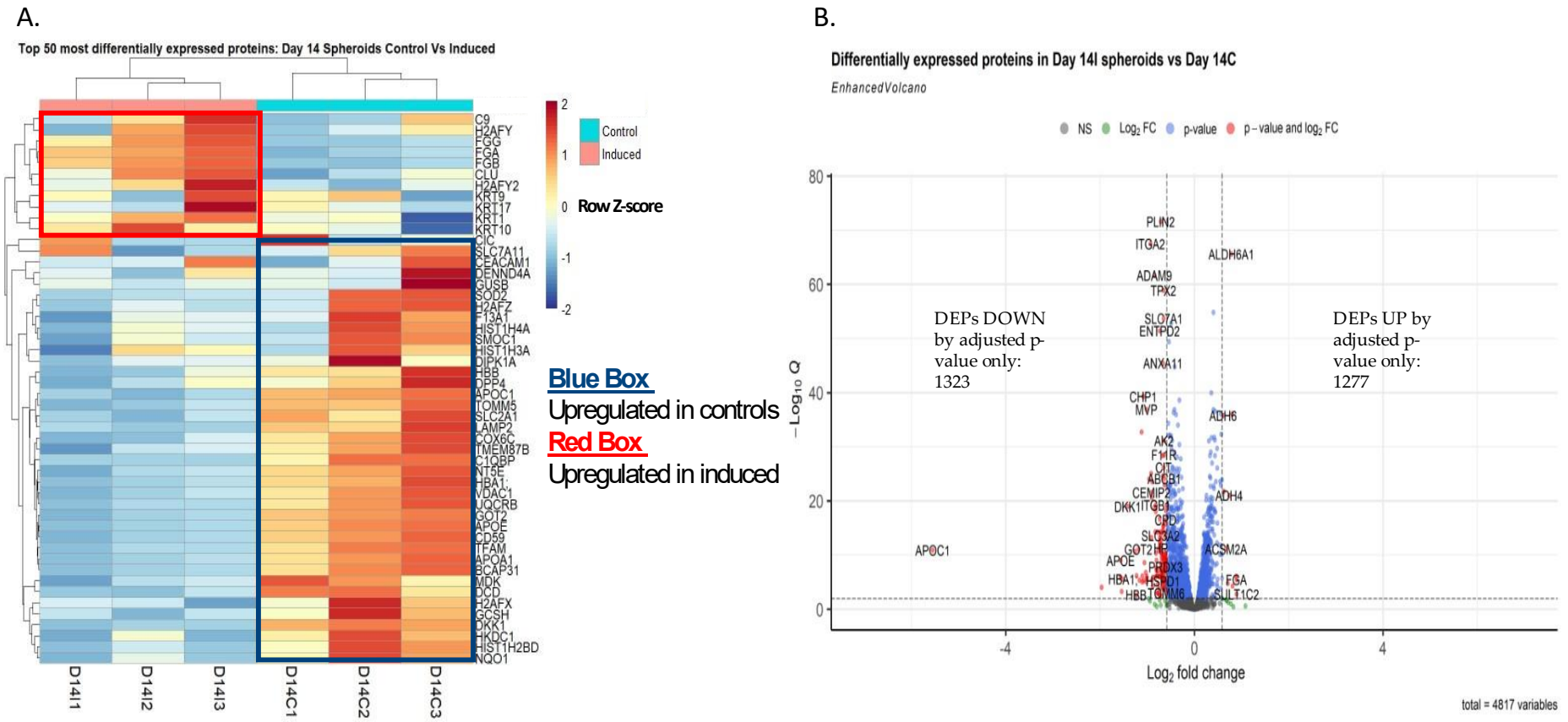
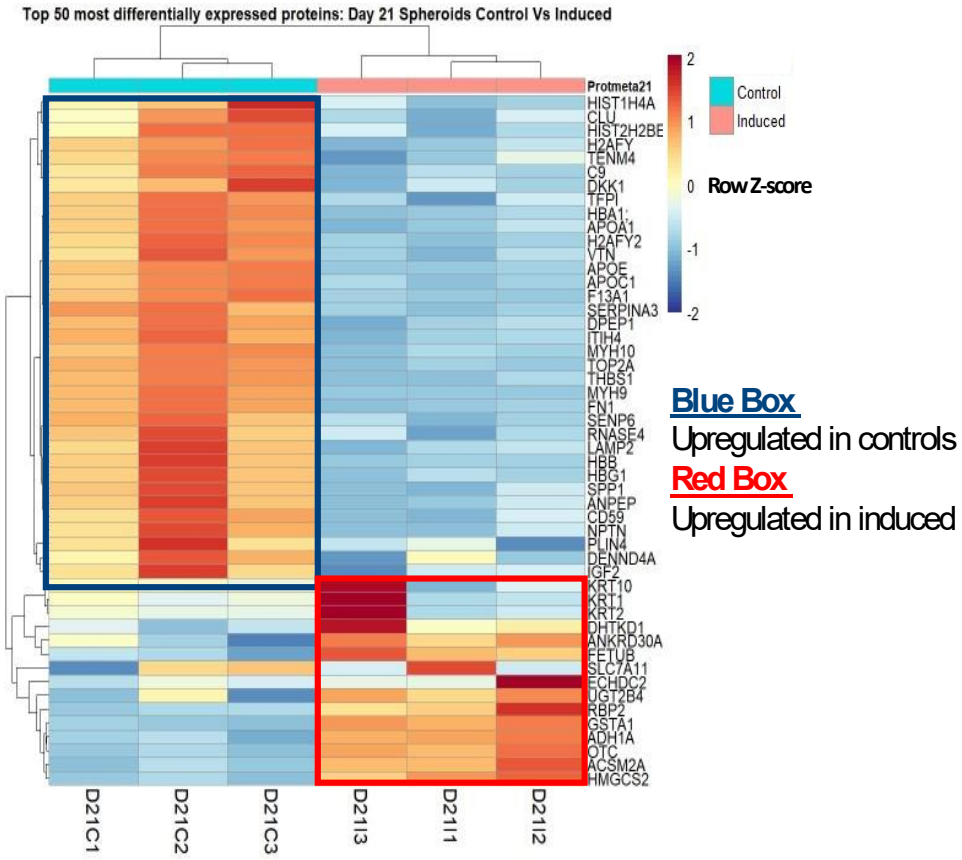


Figure 4.8: Comparisons of differentially expressed proteins between control and induced spheroids from the same time points for A) Paired unsupervised clustering of top 50 most variable proteins groups at day 14 and B). Volcano plot of D14C vs D14I spheroids.

Comparisons of DEPs identified in D21C vs D21I spheroid groups

A.



B.

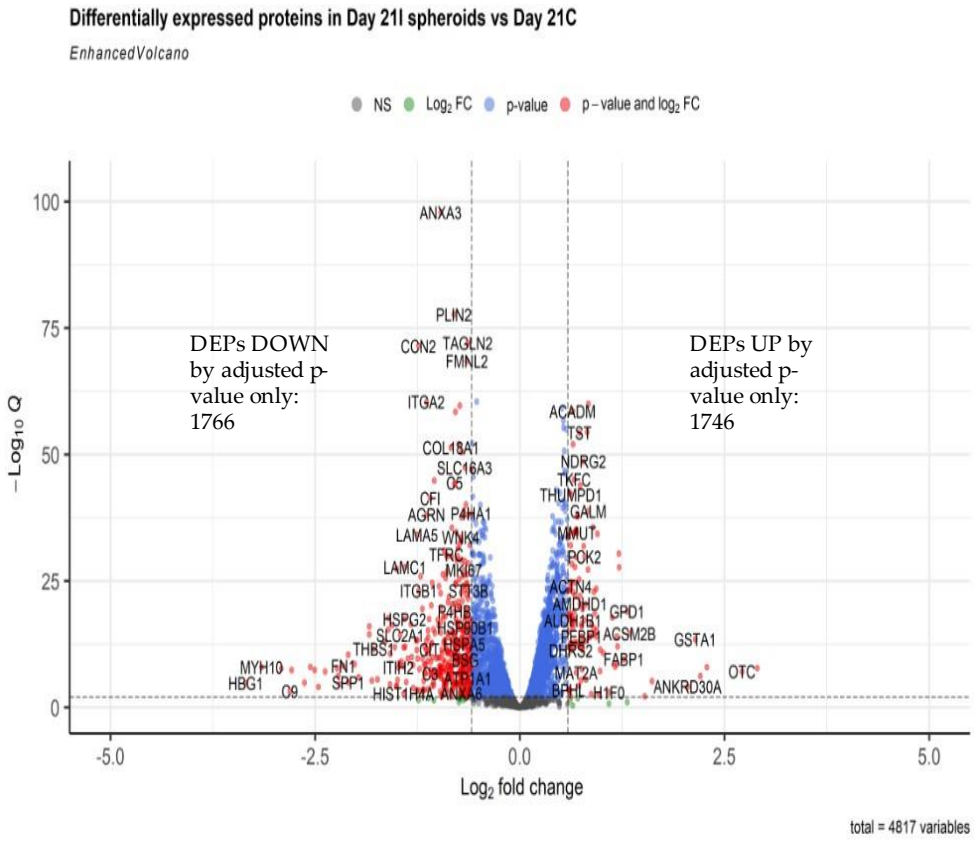
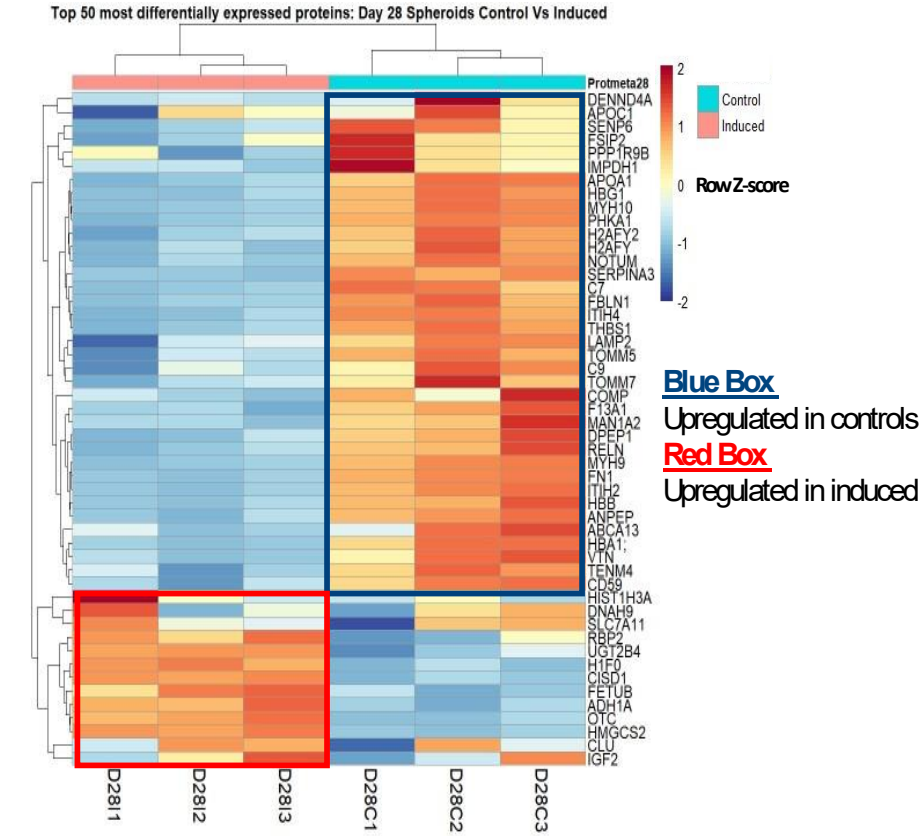


Figure 4.9: Comparisons of differentially expressed proteins between control and induced spheroids from the same time points for A) Paired unsupervised clustering of top 50 most variable proteins groups at day 21 and B). Volcano plot of D21C vs D21I spheroids.

Comparisons of DEPs identified in D28C vs D28I spheroid groups

A.



B.

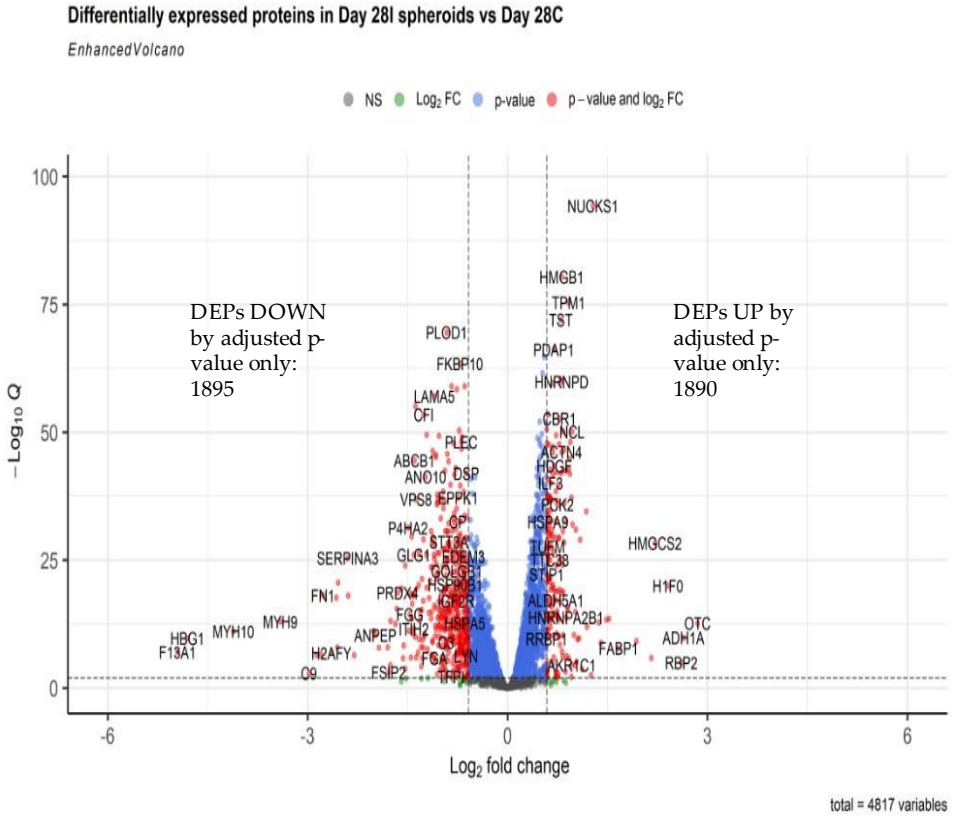


Figure 4.10: Comparisons of differentially expressed proteins between control and induced spheroids from the same time points for A) Paired unsupervised clustering of top 50 most variable proteins groups at day 28 and B). Volcano plot of D28C vs D28I spheroids.

1.1

Figure 4.11:Venn diagram showing the overlap of top 50 variably expressed proteins when comparing spheroid control vs drug treated groups over the time course

4.3.4. Protein Set enrichment analyses of differential expression patterns between control and induced spheroid groups

As is the case for many omics-centric experiments, large scale proteomic experiments produce increasingly large datasets of candidate proteins. Data from these experiments are notoriously challenging to analyse for functional implications without the use of various computational approaches. A typical approach for the analysis of large biological data sets is to investigate the enrichment of gene or protein

sets against various curated datasets. This reduces the high dimensionality of datasets and assigns multiple data features to one or more enrichment term(s) which are associated with a biological pathway, function or structure of interest [287, 288]. Inferences are made using statistical methods, such as hypergeometric testing or Fisher's exact tests, to compare the proportion of genes or proteins within a global or curated list against those present within an associated annotation. Common tools used in this regard are PANTHER [314], GOnet [315], Ontologizer [316], FatiGO [317] and GOrilla [318]. These tools assign a p-value according to the probability that an equal to or greater than enrichment could be observed by chance. This approach has several limitations and inaccuracies as genes and proteins are co-expressed in the majority of organisms. Additionally, with proteomics data, these algorithms consider the proteins matched to an annotation term to be binary in the context of their contribution to the enrichment and does not account for each protein's relative abundance.

While these methods are useful for the identification of pathways or enrichments of interest when the data are curated strategically, as was done in Chapter 3, successful identification often relies on manual investigation which is highly labour intensive and tedious. An alternative approach is to use the popular approach of Gene Set Enrichment Analyses (GSEA) [294, 297, 319, 320] or, in the case of proteomics datasets, Protein Set Enrichment Analysis (PSEA) [321]. Unlike the classical GO enrichment analyses strategies, GSEA and PSEA do not require arbitrary threshold values to assign genes or proteins to a feature but instead, generate ranked lists of all gene or protein ID's based on a feature of interest, such as their relative fold change values. These ranked sets are then assigned a weighted enrichment score (ES) in an annotation set derived from the MSigDB [288], based on the clustering of the pathways matching features on the ranked list. Statistical significance is then assigned using a calculated weighted and normalized enrichment score and a p- or q value is assigned. Enrichment clusters (Top 5 per group) were compared by sub-setting the most significantly enriched pathway term within fold change order lists of significantly

differentially regulated proteins. Enriched terms were plotted as dot plots indicating enrichment terms identified between groups and Cnet plots indicating the proteins involved with each enriched pathway term as well as identifying protein groups shared between enriched terms. The results from the PSEA analyses are presented in Figures 4.12-4.15 below.

PSEA analysis was useful for confirmation of the trends observed between the control and induced spheroids as observed in the previous analyses of the data. Ranked protein lists were enriched and revealed phenotypic differences between control and induced spheroid groups, which became progressively more functionally adapted over the time course. Indeed, DEPs identified within control spheroids were associated with spheroid growth and architecture, as demonstrated by the enriched terms exogenous protein binding, integrin binding, collagen binding and carbohydrate binding. Conversely, the terms enriched from upregulated proteins in drug treated groups were related to enzymatically driven processes (oxidoreductase activity), suggestive of functional changes in the cell. Even though exposed to a drug cocktail constantly for 14 to 28 days in these induced groups and a significant increase in general metabolic enzyme expression is noted, the repertoire of clinically relevant CYP450 drug metabolising mono-oxygenase enzymes were still essentially absent from the observed proteome, even prior to filtering of datasets. This was despite the capacity for induction of CYP 450 enzymes, at gene transcription level in HepG2 cells, which has previously been reported by others [227, 262, 272, 304]. Interestingly, the CYP450 enzymes have not been detected in the proteome at significant levels in cultured HepG2 cells at any point during this research project.

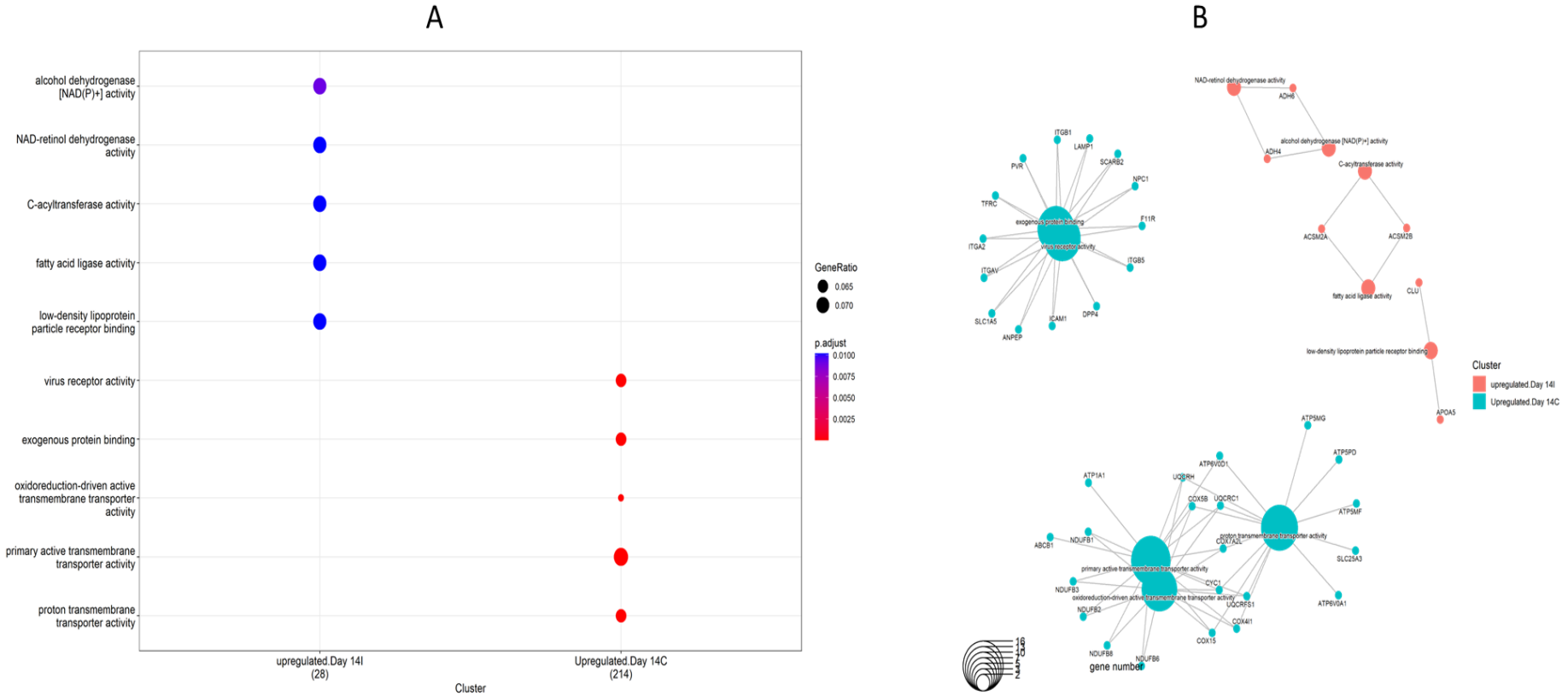


Figure 4.12: Protein set enrichment analyses of differentially enriched proteins between control and induced spheroid groups at Day 14 in culture. A) Dot plot indicating terms significantly enriched between groups, coloured according to adjusted p-value. B) Cnetplot indicating proteins associated with respective enriched terms

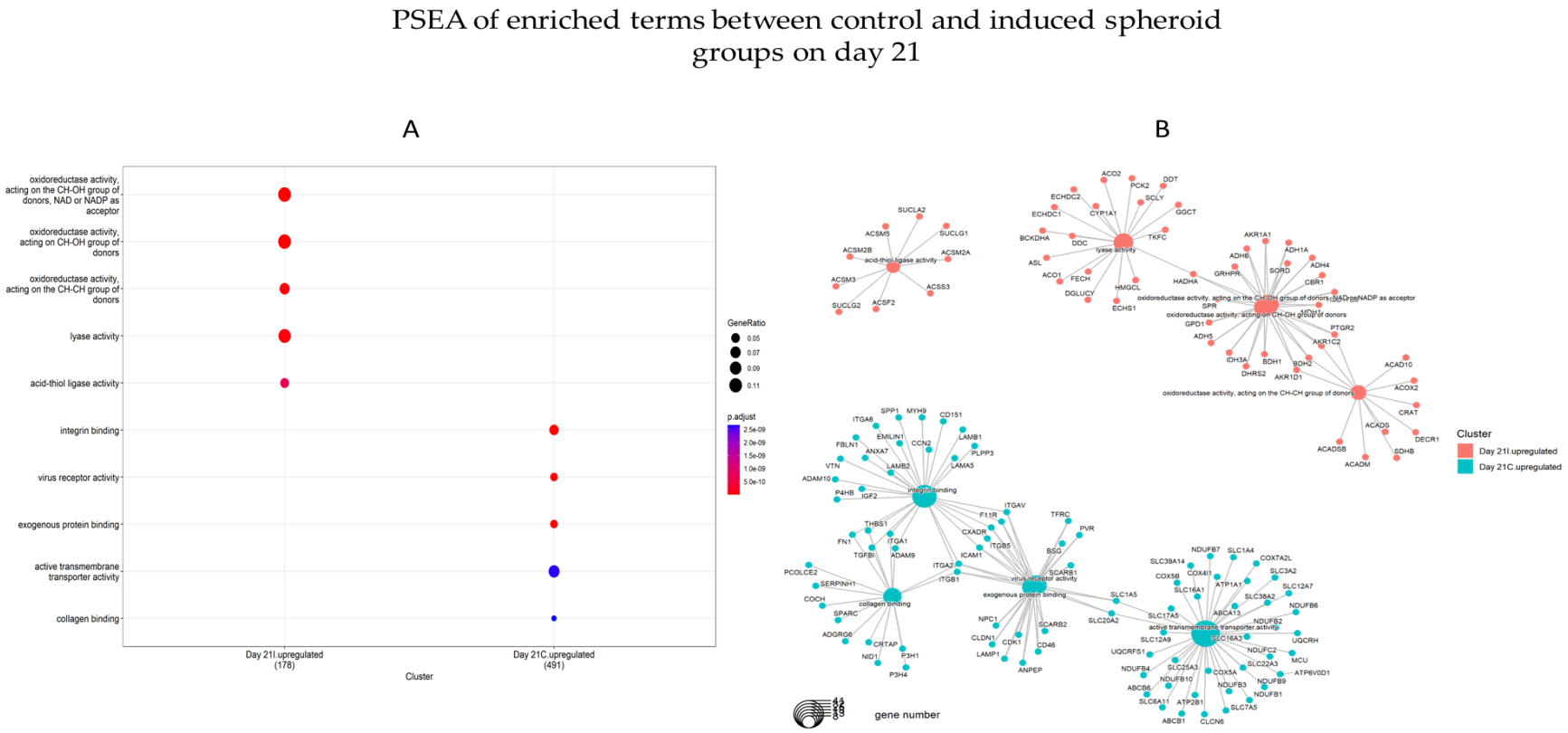


Figure 4.13: Protein set enrichment analyses of differentially enriched proteins between control and induced spheroid groups at Day 21 in culture. A) Dot plot indicating terms significantly enriched between groups, coloured according to adjusted p-value. B) Cnetplot indicating proteins associated with respective enriched terms

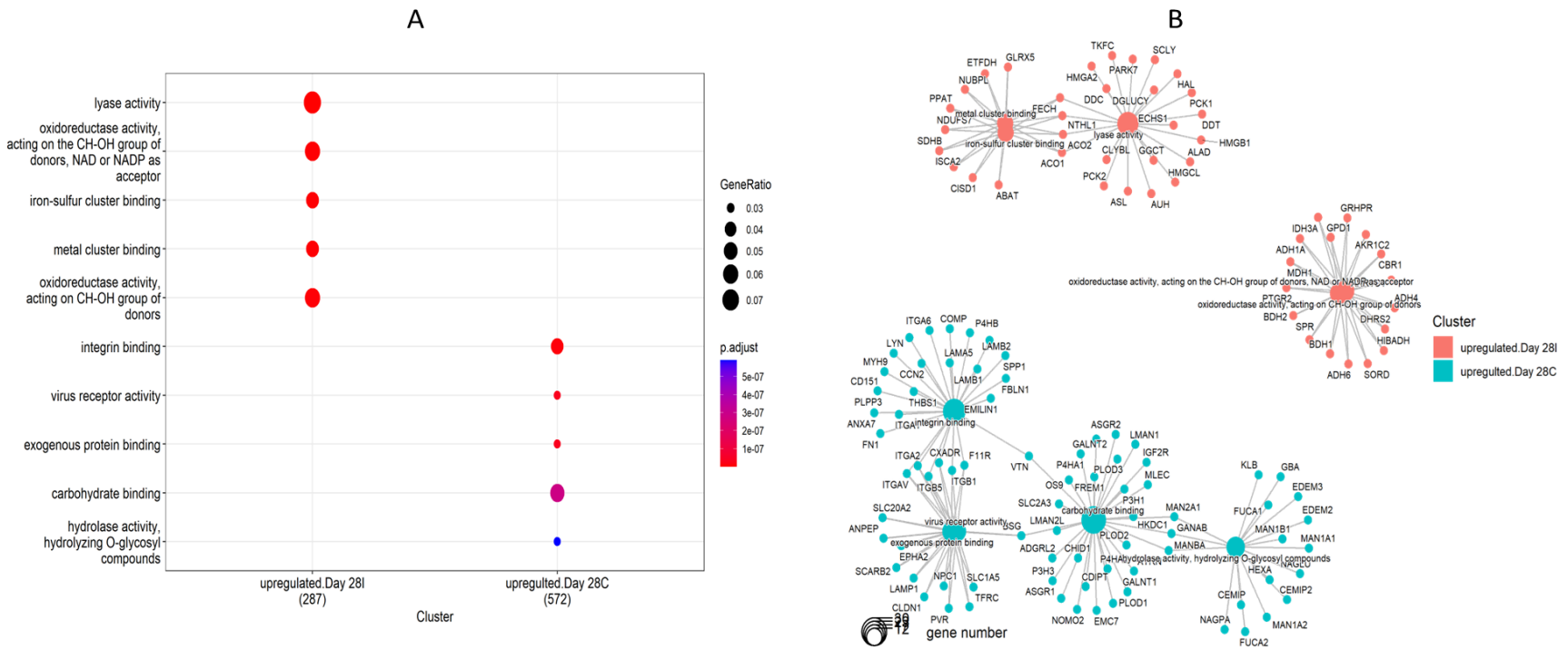


Figure 4.14: Protein set enrichment analyses of differentially enriched proteins between control and induced spheroid groups at Day 28 in culture. A) Dot plot indicating terms significantly enriched between groups, coloured according to adjusted p-value. B) Cnetplot indicating proteins associated with respective enriched terms

The dissimilar expression profiles of control verses induced spheroids cultured for the same time course can be further examined using a more targeted analysis. For example, HepG2 spheroids without drug treatment was previously shown to progressively upregulate protein involved in the extracellular matrix [182]. This was not observed to the same extent in the case of drug induced spheroid groups. To investigate this in more detail, another useful analyses method was used, pathway analysis [198, 289, 296, 320]. The expression of proteins involved in the extracellular matrix were compared using the PathView package in R and searched against the KEGG pathway term for ECM-Receptor Interaction (hsa04512). In Figure 4.9, comparisons are made between control and induced spheroids, as done previously where higher fold changes in favour of control groups are expressed with negative values and positive fold changes are representative of upregulated in induced groups. In general, higher expression of ECM-Receptor binding proteins persist most strongly in the latter part of the culture time-course, however when comparing spheroid groups directly, control spheroids show a consistently higher expression of most ECM proteins which is amplified in later timepoints in culture. It should be noted that, relative to monolayer groups, both control and drug treated spheroids upregulate ECM proteins, but the extent of these changes differs.

Progressive downregulation of ECM -Receptor binding D14I vs D14C groups

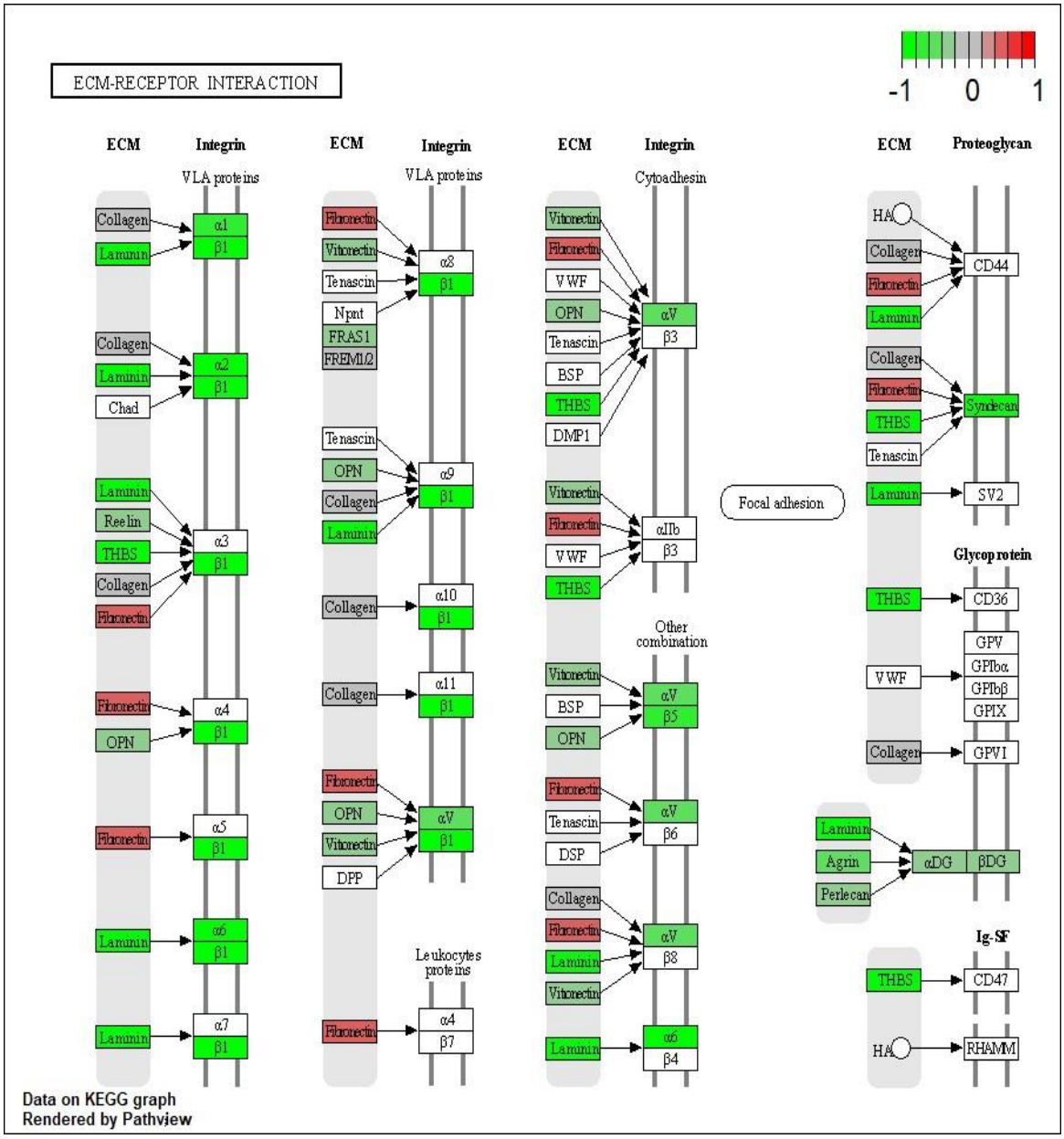


Figure 4.15: Pathway analysis for ECM-Receptor proteins for D14C vs D14I. Convention: red = higher in I groups, green = higher in C groups. Grey = no change and white blocks = protein missing from data

Progressive downregulation of ECM-Receptor binding D21I vs D21C groups

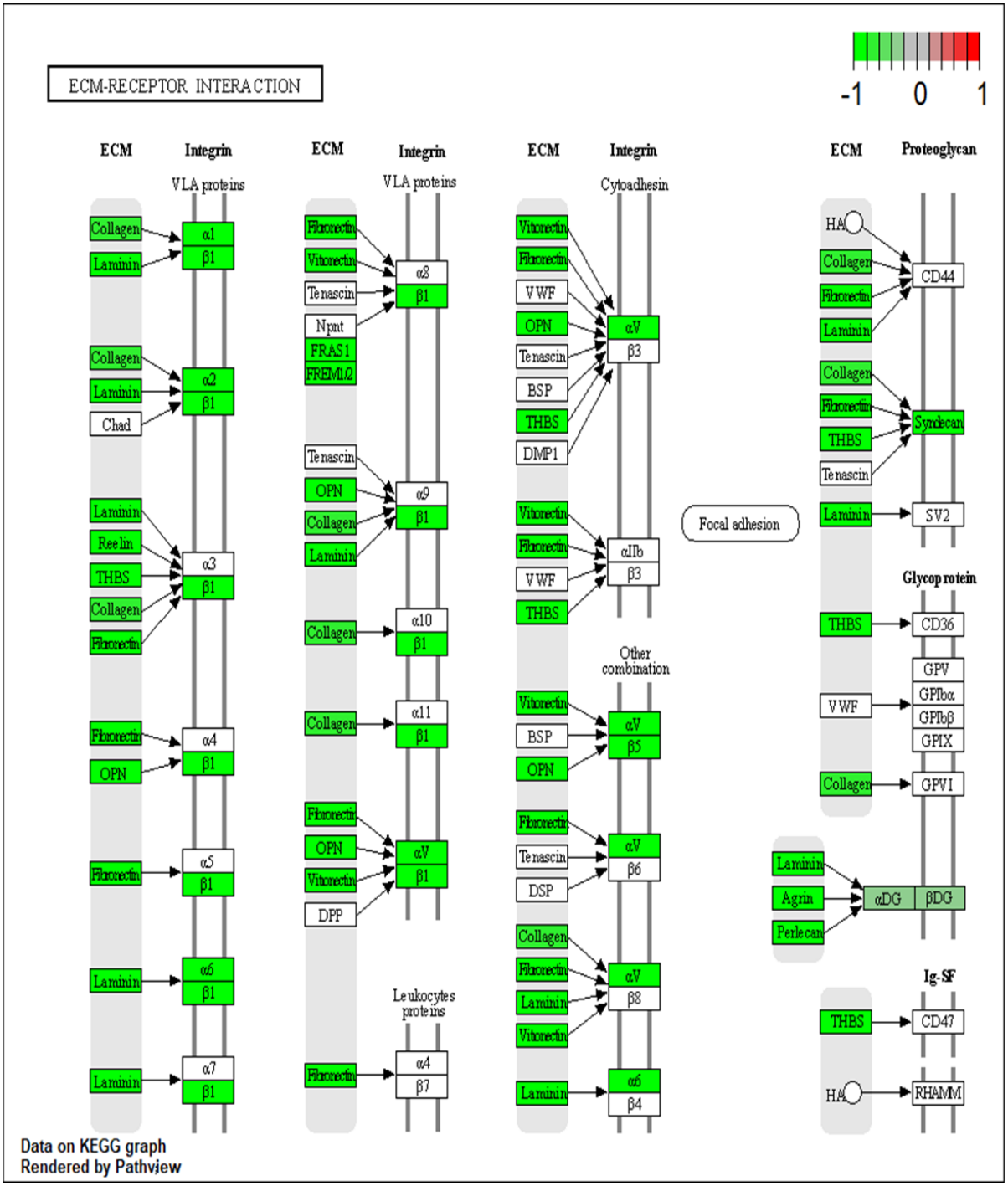


Figure 4.16: Pathway analysis for ECM-Receptor proteins for D21C vs D21I. Convention: red = higher in I groups, green = higher in C groups. Grey = no change and white blocks = protein missing from data

Progressive downregulation of ECM-Receptor binding D28I vs D28C groups

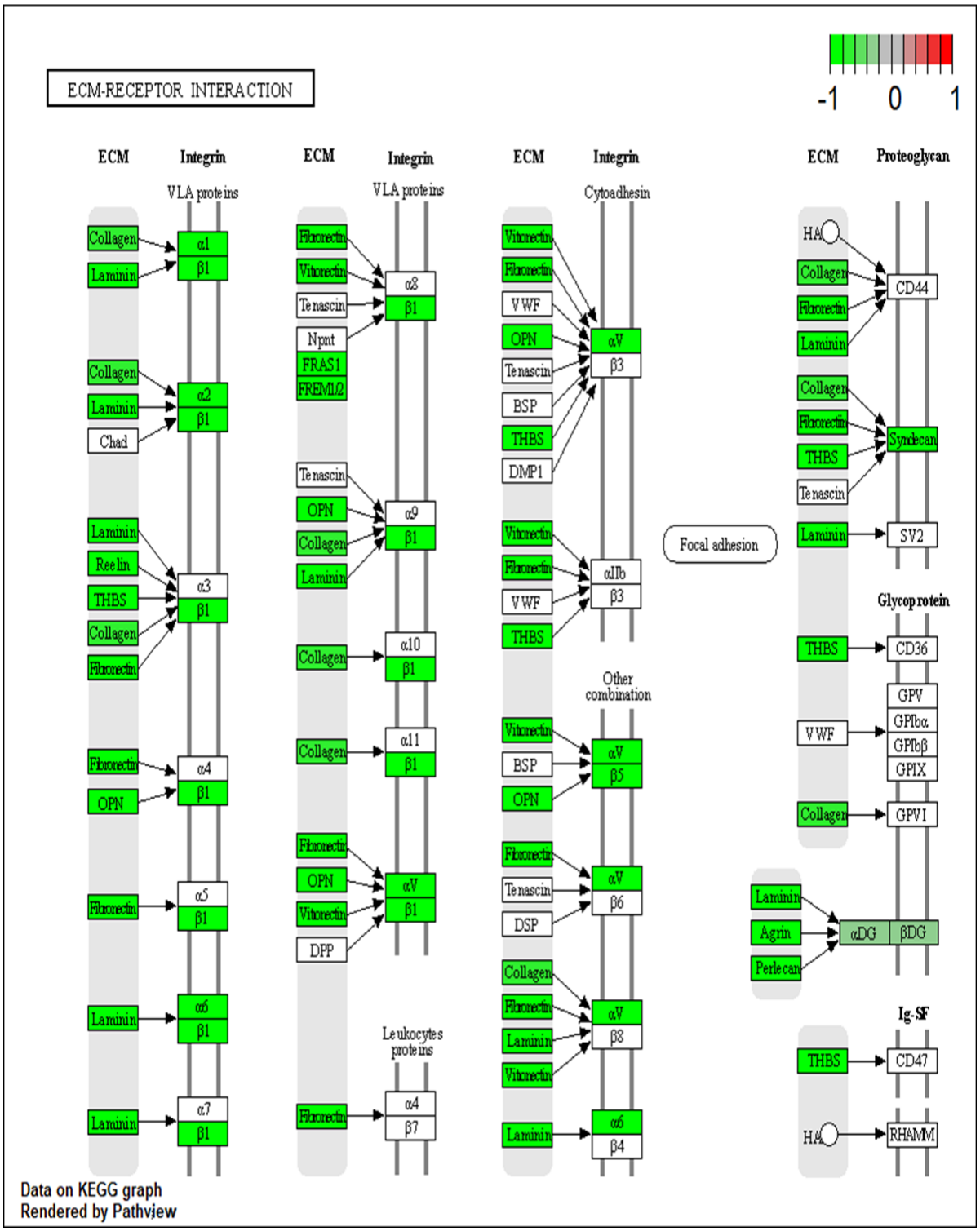


Figure 4.17: Pathway analysis for ECM-Receptor proteins for D28C vs D28I. Convention: red = higher in I groups, green = higher in C groups. Grey = no change and white blocks = protein missing from data. Immunohistochemistry identifies evidence for differential nuclear localisation of transcriptional regulator HNF4α

Successful clearing and immunohistochemistry labelled confocal microscopy resulted in spheroids with a normalised refractive index and greater laser penetration into deeper cell layers. Figure 4.18 shows a representative image for a HepG2 spheroid before and after clearing using the FOCM method [301]. Minimal distortions were seen with regards to overall spheroid shape and structure sparing a small haze which is observable surrounding the FOCM cleared spheroid, not present around the uncleared. This is possibly due to dissociating cells resulting from either mechanical damage due to handling or degradation due to exposure to the clearing reagent.

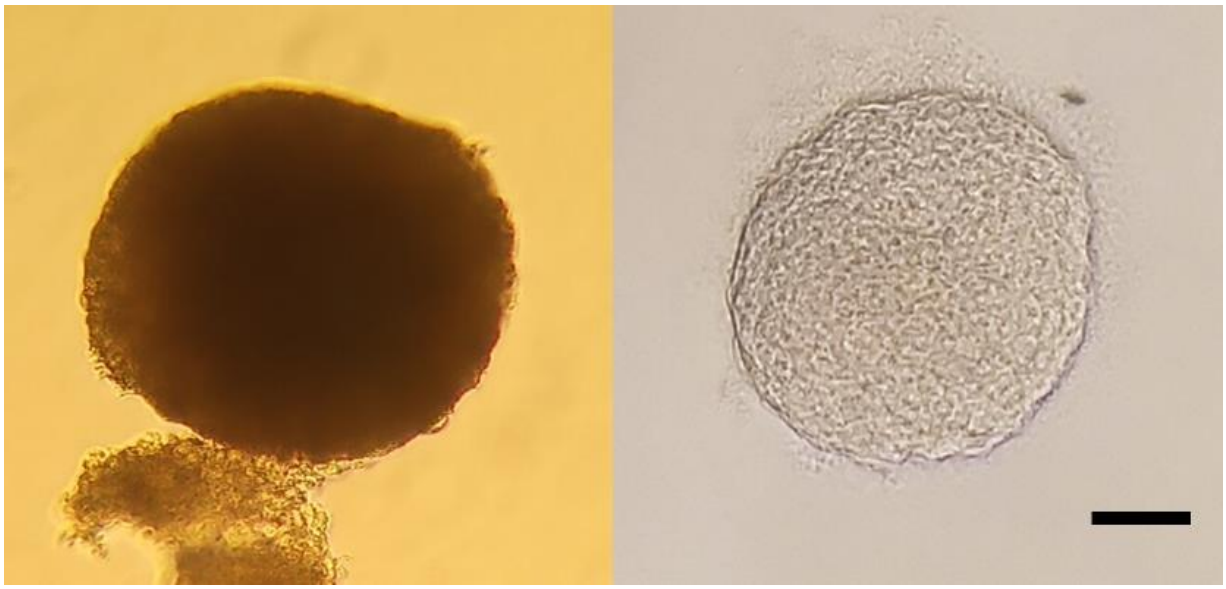


Figure 4.18: Uncleared HepG2 spheroid (left), HepG2 spheroid cleared using the FOCM method (right). Scale bar = 100 μm.

As discussed in previous sections, HNF4α is a transcriptional regulator which is required for both hepatic lineage specification and maturation. Post-developmental HNF4α is expressed within mature hepatocytes for maintenance of function. While the expression of HNF4α was shown to be upregulated in comparison to monolayer cultures but maintained relatively stable throughout spheroid cultures here (Figure 4.19) we are able to see its differential abundances with regards to nuclear localisation with seemingly higher nuclear localisation seen in induced spheroid groups.

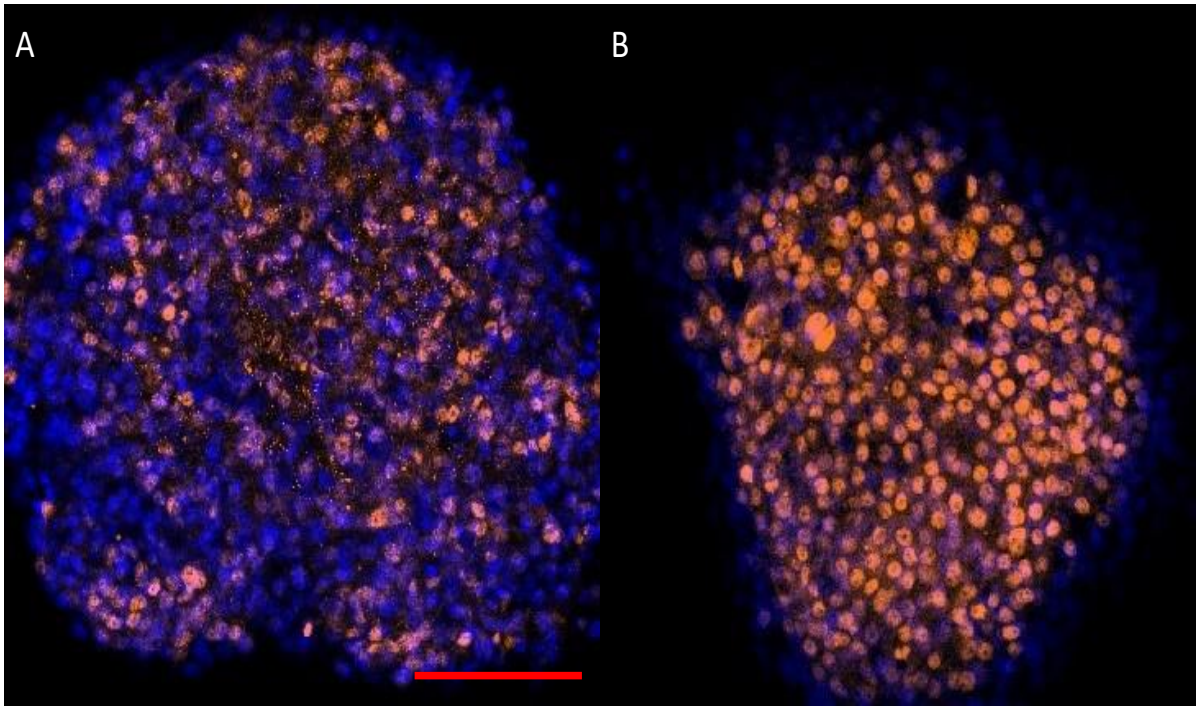


Figure 4.19: Confocal images of A) Control spheroid at day 28 in culture and B) Drug treated spheroid at day 28 in culture. Both immunohistochemically labelled with HNF4α (orange) and DAPI (blue). Scale bar= 100 μm

These changes were quantifiable by measuring fluorescent signal from captured images using the "Measure" tool built into the Zeiss Zen Blue software package. Relative fluorescent intensities were recorded and normalised to DAPI as done previously in Chapter 3 with the analysis expanded to include comparisons to drug treated spheroid groups. Figure 4.20 indicates progressive expression of these marker proteins as recorded by confocal microscopy analyses.

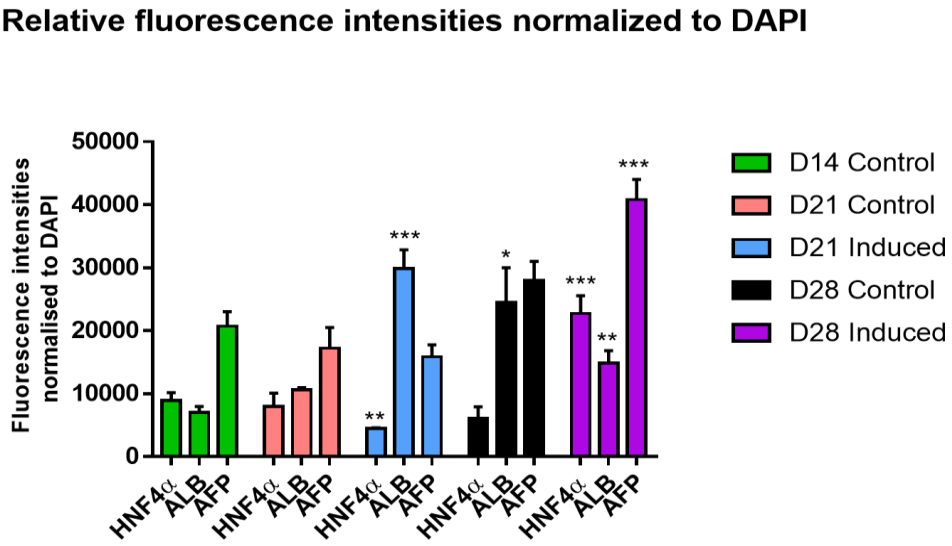


*Figure 4.20: Relative quantitation of hepatic marker protein expression in control and induced spheroid groups. Error bars as mean ± SEM, significance relative to day 14 control as * for $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Determined according to multiple comparisons t-test with 1% FDR correction, n=5-8.*

Though, in general the protein expression trends recorded using semi-quantitative confocal microscopy analyses matched those observed within curated proteomic datasets, the pattern of expression for HNF4α did not recapture the significant increase seen within D28 induced groups here. This may be due to several reasons. The first is that, since HNF4α is predominantly localised to the intra-nuclear compartments where it is able to exact its biological functions, it may have been lost during pelleting of cell membranes during the sample collection phases of proteomic sample preparation. The second is that the observed changes in expression are due an experimental artifact.

To further investigate these theories, we employed PSEA to monitor the relative normalised enrichment scores (NES) for protein sets identified as downstream targets for HNF4α between the similar pairing of D28C vs D28I spheroid groups. These signature sets were obtained from MSigDB in transcriptomic experiments conducted by others[322]. Figure 4.21 indicates the relative NES for downstream HNF4α targets.

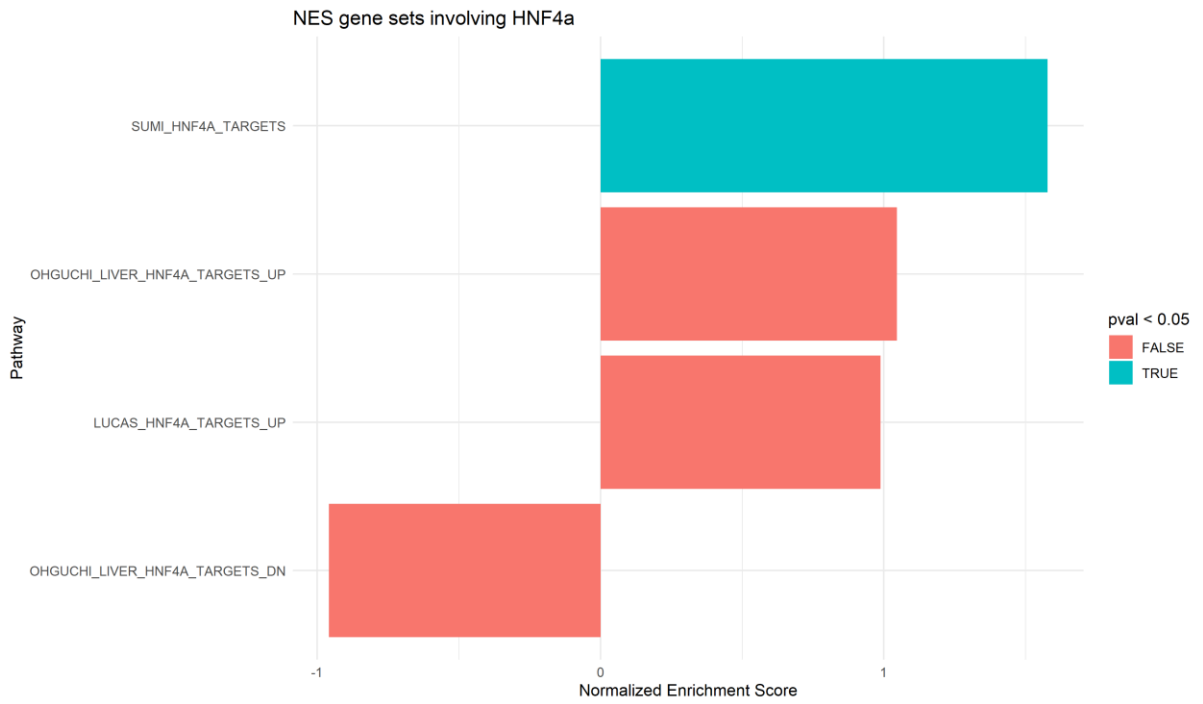


Figure 4.21: PSEA generated NES for differentially expressed protein groups matching gene sets of downstream targets of HNF4α in D28I vs D28C spheroids. Positive value for NES is indicative of enrichment towards induced groups, negative NES is enriched in favour of controls

While clear trends were observed for targets upregulated downstream of HNF4α activation, PSEA enrichment for ¾ of the pathways were not able to reach statistical significance. This is possibly due to coverage differences between transcriptomics and proteomics datasets where proteomics is not able to cover gene set enrichments with enough protein:gene matches to confer statistically significant enrichment. However, in keeping with the central dogma of molecular biology, it is reasonable to predict that transcription and protein expression are at least somewhat correlated. To further visualise the usefulness of this method. Proteins which were successfully identified within curated proteomics data that were matched to the most significant PSEA enrichment (SUMI_HNF4A_TARGETS). In total 11 proteins were able to be matched to the pathway gene dataset (11 out of 30). The DESeq2 log2 foldchanges for the identified proteins are presented in Table 4.6 below.

Table 4.6:List of proteins identified against signature gene set for HNF4α activation

Gene name	Uniprot Accession	log2FoldChange	lfcSE	stat	padj
ACSM2B	Q68CK6	1.014	0.123	8.222	0.000
AKR1D1	P51857	0.450	0.068	6.584	0.000
BDH1	Q02338	0.881	0.063	14.086	0.000
BHMT	Q93088	1.401	0.225	6.239	0.000
FABP1	P07148	1.666	0.286	5.836	0.000
KANK4	Q5T7N3	-0.173	0.052	-3.347	0.001
LGALS2	P05162	1.023	0.085	11.998	0.000
PLA2G12B	Q9BX93	0.211	0.095	2.232	0.035
RBP2	P50120	2.608	0.570	4.578	0.000
SLC13A5	O75897	-0.240	0.063	-3.819	0.000
SULT1C2	O00338	0.193	0.090	2.147	0.043

4.3.5. Evaluation of the capacity for drug metabolism in HepG2 spheroids

The proteomic transitions and potential benefits conferred on HepG2 cells, by virtue of extended three-dimensional culture, has been explored extensively above and in Chapter 3. However, despite proteome changes apparently indicating trends toward a cell status more representative of the *in vivo* tissue as a result of the culture process, these are arguably meaningful in the context of a hepatocyte model, if their principal utility relating to xenobiotic metabolism does not mimic the *in vivo* reality. To assess metabolic competence, an LC-MS/MS based metabolomic study using a targeted MRM based method and a validated phenotyping drug cocktail was performed. Figure 4.10 shows an extracted ion chromatogram (XIC) demonstrating the separation of probe analytes achieved using the method. Figure 4.11 is a representative calibration curve demonstrating method linearity for one of the probe analytes (midazolam).

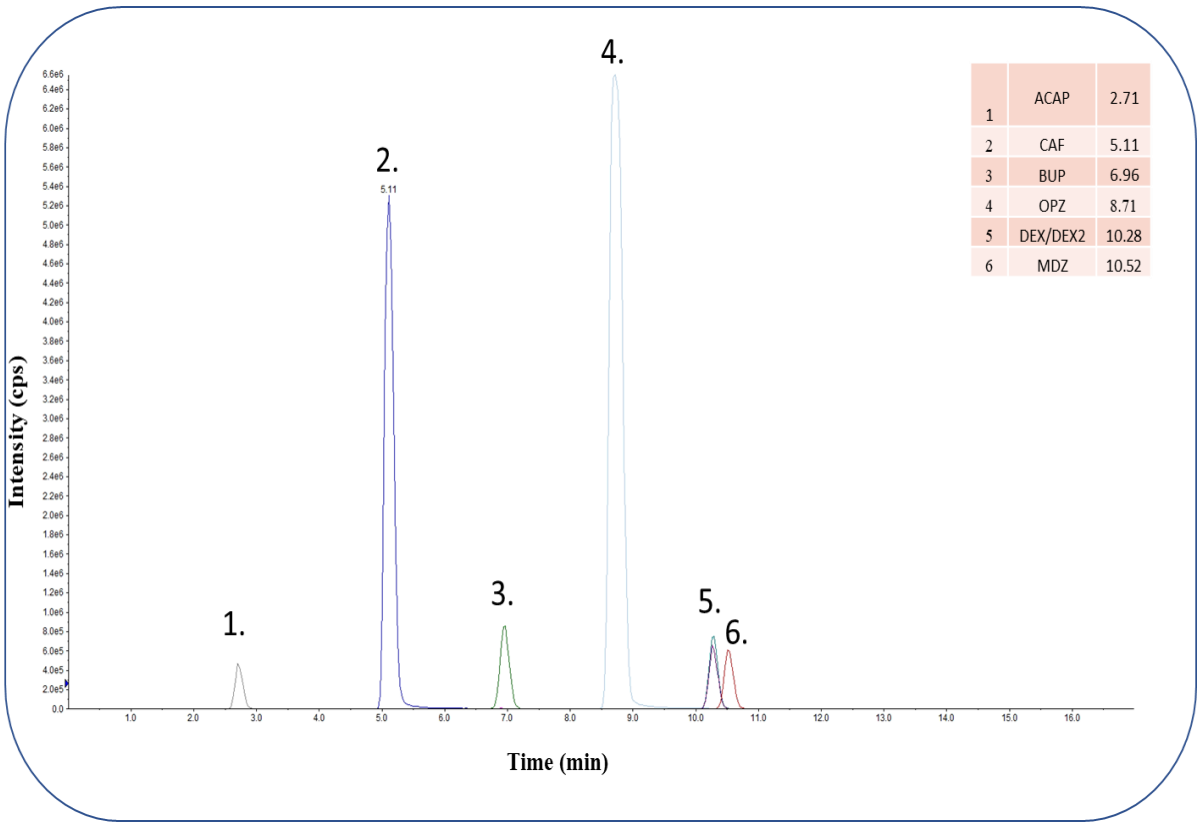


Figure 4.22: Representative XIC for separation of probe drug analytes

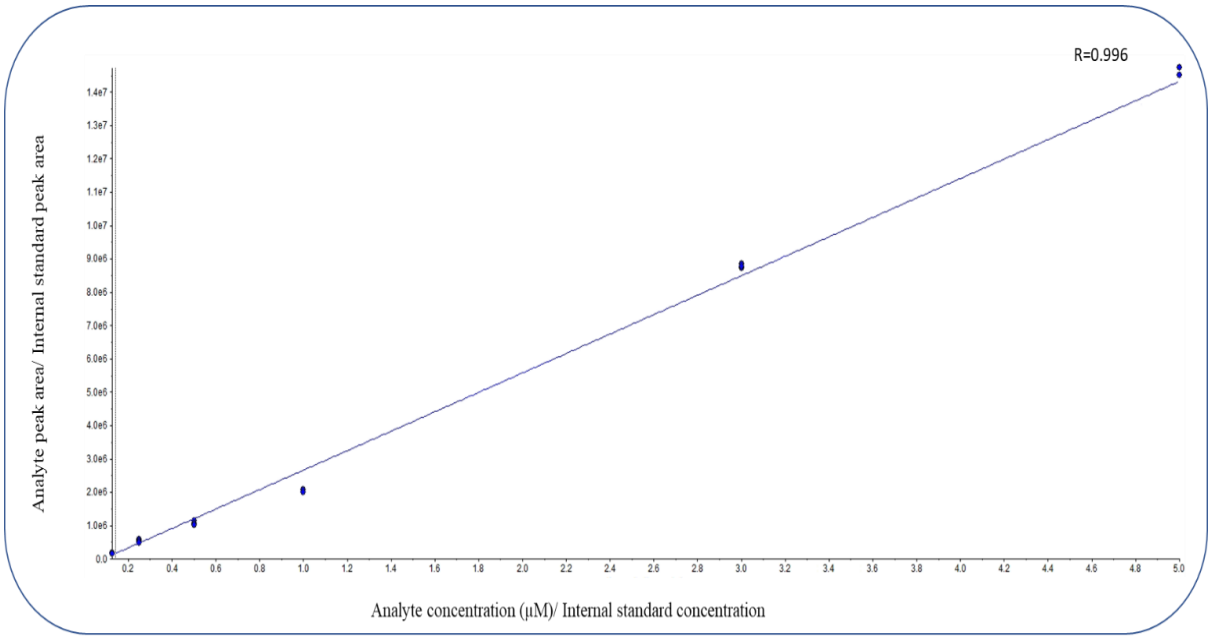


Figure 4.23: Representative calibration curve for Midazolam (MDZ)

The capacity for the liver to upregulate metabolic pathways in response to drug exposure is well established. HepG2 spheroids, when investigated for their ability to recapitulate this response, have shown variable results. Some studies report no detectable metabolic activity at all, while others infer this activity from upregulation of gene expression [227]. The metabolomic profiling of HepG2 spheroid cultures in this study gave mixed results with some probe drugs showing expected metabolic products at high enough concentrations to calculate a metabolite-parent drug ratio, while others did not produce any detectable expected metabolic products. This is especially true in the cases of drugs with a tendency to be spontaneously chemically modified such as the conversion of caffeine to paraxanthine (facilitated by CYP1A2) resulting in any potential metabolic formation of paraxanthine to not be resolved against the background (data not shown). In the case of flurbiprofen (CYP2C9) reduction of ionisation efficacy may have been a contributing factor causing a relatively lower overall sensitivity due to negative mode ionisation. This is in contrast to bupropion (CYP2B6) which despite having good sensitivity and low background was not detectable at all. Figure 4.12 shows the observed conversion of probe drug to the expected metabolite, expressed as a ratio of metabolite to parent drug. Overall, the conversion of parental drug to metabolite ratios were low. While some probes had a consistent increase in the metabolite-parent compound ratio for induced versus control spheroids, other probes produced inconsistent high metabolite-probe drug ratios in either control or induced spheroid groups over time. These results could be due to the fact that the initial concentration of the probe drug may have been at a concentration that was excessive when considering the total protein present in the spheroid samples and the flux rate across the cell membranes. Parent drugs which did not show a detectable metabolic conversion resulting in a metabolite which was quantifiable against the background were not shown here. In general, metabolite-probe drug ratios were highest in the Day 21 induced spheroid groups compared to other culture time points. Conversion of omeprazole (CYP2C19) was ambiguous and

did not exclusively favour the induced spheroid groups, with the Day 28 group showing increased metabolic activity within control spheroids. Metabolism of midazolam (CYP3A4) showed the most robust increases in induced groups and showed the highest relative difference at Day 21. Metabolic conversion of dextromethorphan (CYP2D6) showed a similar trend to that of midazolam but was not statistically significant except for the 3 h incubation of the Day 21 induced spheroid.

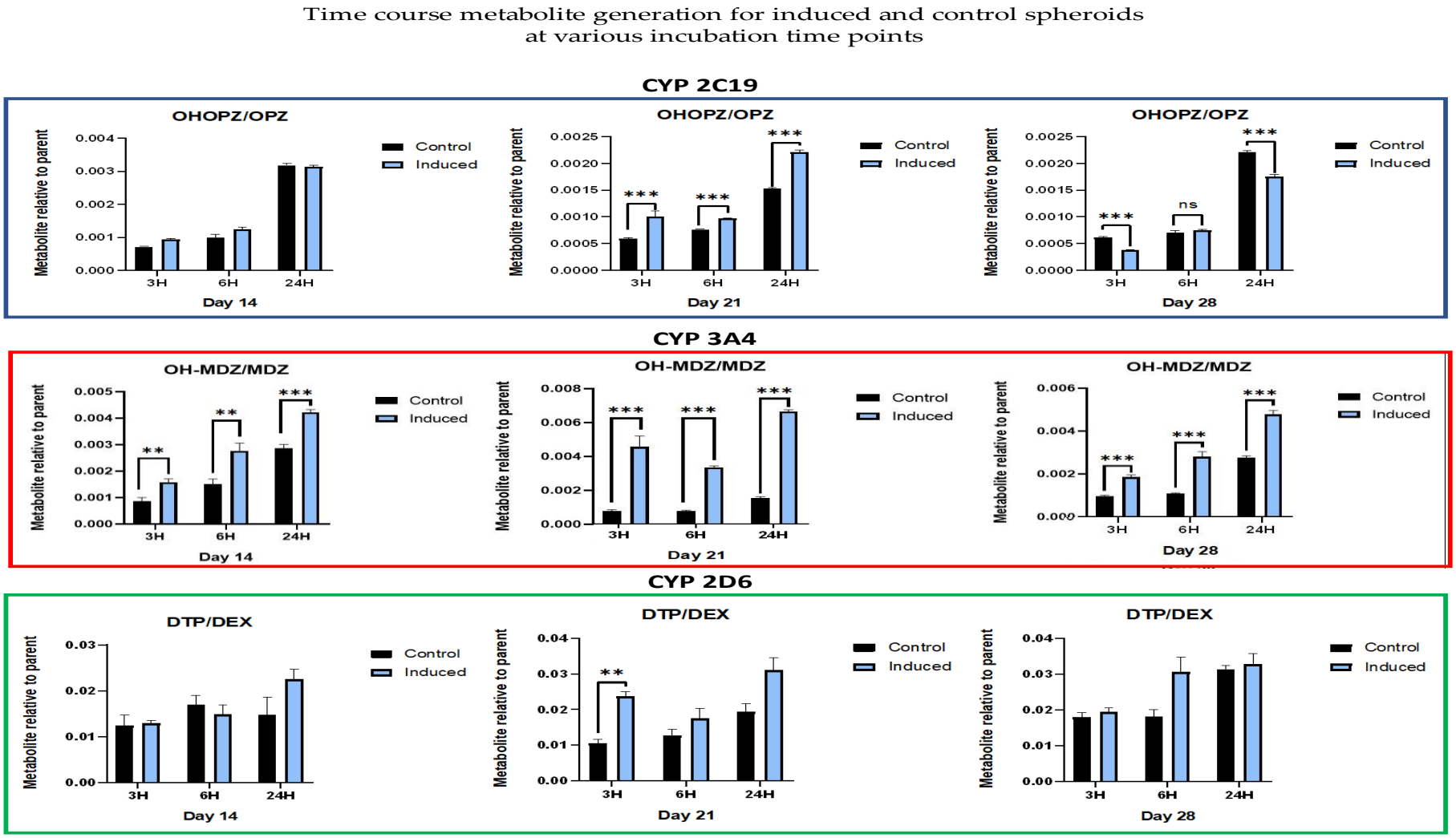


*Figure 4.24: Probe drug to metabolite conversion by induced and control spheroid groups by time. Error bars as mean ± SEM, significance relative to day 14 control as * for $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Determined according to multiple comparisons t-test with 1% FDR correction.*

4.4. Conclusion

Quantitative proteomics, using isobaric tagging, was successful for the identification and quantitation of over 4800 proteins in biological replicates across the 7 experimental conditions surveyed in this study. As reported above (Chapter 3), the major changes in the spheroid proteome could be linked to the time in culture, and there appears to be a shift to a phenotype that is less orientated towards proliferation and shows more interaction with the extracellular matrix and potentially metabolic functionality. This part of the study expanded on the concept that the proteomic changes were related to time course that the spheroids were in culture and was investigated alongside the effect of long-term exposure to an induction drug cocktail added into the culture media. Perturbations to the proteome differed between the control and induction conditions compared to both the initial monolayers from which the spheroids were grown and to equivalent spheroids from the same paired timepoint. A potential mechanism by which drug exposed spheroid groups further change protein expression profiles is by modulation of HNF4α transcriptional activity as shown by the differential expression of downstream elements of HNF4α activation which were evident even at a protein level

The adoption of cross-omic bioinformatics analysis methods have proven useful in the elucidation of the underlying mechanisms for the proteomic changes observed and have further served to characterize these changes in an objective and simplified manner. The differential expression of proteins that are involved in metabolic pathways in drug treated spheroid groups, was supported by the increase in drug metabolism seen within these groups. Though these differences are undoubtedly present, without a representative control for in vivo counterparts, the value of these changes is difficult to validate, and their relative representativeness is likely still insufficient for accurate recapitulation of hepatocyte xenobiotic metabolism.

Chapter 5. Global discussion and acknowledgment of limitations

5.1. Overview and contextualisation

Scientifically robust, clinically representative, and functionally translatable preclinical models have become increasingly required within pharmaceutical research. Regulatory bodies are constantly adapting the minimum requirements for approval and providing new mandates to the pharmaceutical industry. Mandates such as the FDA [323] requiring biotechnology companies to rethink and revisit preclinical studies and clinical trial designs. This shift occurred naturally though the industry and markets progressive adoption of targeted therapies for oncology indications. Historically, cytotoxic oncology agents were granted approvals utilising the widely adopted 3+3 dose escalation study design geared to find the maximum tolerated dose (MTD) that in the case of cytotoxic agents is also conventionally the most efficacious dose. However, in the context of targeted agents, the MTD may be well above the minimally efficacious dose as successful treatment is more reliant on receptor occupancy rather than attaining the highest dose possible. Consequently, this prompted a shift from a MTD informed approach to a model informed approach that is based on both pharmacokinetic and pharmacodynamic endpoints, which ultimately spares patients needless toxicity. Similarly, in the fields of preclinical modelling researchers are becoming more informed on the appropriate utilisation of biological models and realising the limitations of models currently in use. Preclinical models need to pivot from simply matching a cell type to a model to addressing and characterising function related the underlying biological mechanisms and questions at hand.

This study set out to answer several questions related to better characterising HepG2 spheroid cultures within the context of preclinical hepatic models for drug screening and metabolomic applications. Firstly, proteomic implications of spheroidal cell culture on HepG2 cells were characterised. Secondly, the temporal aspect of spheroid culture was

assessed to determine whether culture time further contribute to these altered proteome states. The analyses were then expanded to establish if these potential modulations could be manipulated towards more metabolically active states, as in the case of the metabolic adaptations which occur *in vivo.* Lastly, to measure if the proteomic shifts observed within both control and drug induced cultures were functionally translatable as clinically relevant.

5.2. The use of quantitative proteomics to determine and characterise phenotypic changes

Generation of robust biological data can be innately confounded by selection of inappropriate quantitative methodologies for endpoint assessments. Similarly, technical and processing workflows can introduce intrinsic biases. In the case of the labelled proteomics experiments conducted within this study, TMT labelling and MS based proteomics were able to, with reasonably high coverage and reproducibility, identify and show relative quantitative global proteomic changes undergone by HepG2 cultured as spheroids in the context of answering biologically relevant questions as highlighted above. However, it should be noted that workflow limitations were still present. Sample collection and lysis may bias the proteomic coverage which would be transferred to all downstream assays. When extracting proteins into aqueous solutions, lipophilic and other water-insoluble proteins may be discarded or pelleted during the removal of debris. These biases may be further compounded during the multiple sample transfer and clean-up steps involved in the TMT labelling protocol. The use of high pH reverse phase chromatography using C18 stationary phase, while being widely adopted and appropriate for many proteomic applications, may result in the retention of certain protein groups [324], several of which could be of high interest in the context of this study, like the predominantly membrane bound CYP450 enzymes [325]. Proteins of interest may have been lost during the fractionation and recombination or during solid phase

extraction, where both strategies utilised reverse phase C18 elution. Additional sources of bias may be from large changes in pH causing peptide degradation or precipitation, loss of ionisation capacity or poor tryptic cleavages/miss-cleavages due to detergent or lipid binding.

Capacity for identification of highly resolved biological pathways is greatly dependent on instrument capabilities.

Orbitrap mass spectrometers are known for their high mass resolution, accuracy, and sensitivity, making them useful for a wide range of applications in fields such as proteomics, metabolomics, environmental analysis, and drug discovery. They are also used in imaging mass spectrometry, which allows researchers to map the distribution of molecules in a sample. Orbitrap mass spectrometers make use of Higher-energy Collisional Dissociation (HCD) which is useful when paired with TMT labelled samples. The HCD allows for more robust fragmentation of reporter tags which is essential for high resolution and robust relative quantitation.

Selection of analysis software and search engines for protein identification assignments are also important. The use of differing methods for sample processing may yield different endpoint datasets which, when not carefully validated can lead to study limitations.

In this study, the differences between two mainstream stochastic models for differential expression profiling were employed. While the current industry convention supports the use of multiple comparison t-tests with paired FDR corrections other methodologies such DESeq2s population size factor normalized negative binomial generalized linear model for the identification of differential expression may serve useful, especially when introducing FDR corrections, such as the Benjamini-Hochberg method. DESeq generated dataset objects may also be directly inputted into established downstream analyses pipelines with a greater number of comparative datasets than those making use of strictly proteomic based reference data. Naturally, new methodologies cannot be adopted

without considering potential limitations, such as the disparity in coverages across transcriptomic and proteomic datasets, although with the progressive improvement of mass spectrometers, sample preparation techniques and bioinformatics pipelines, these differences may become progressively smaller as proteomics technologies evolve.

5.3. The use of alternative methods for measuring protein expression

Throughout the sciences, there are very few instances where a single form of analysis is universally suitable especially for large data sets. The same can be said for proteomic analyses, where the adoption of multiple endpoint experiments may prove useful for the appropriate characterisation of proteomic signatures. While MS based proteomics is strongly preferred for large scale quantitation of differential protein expression, simpler methods, such as immunohistochemistry, may be employed when there are fewer measurable endpoints or a targeted approach is required, or where application of conventional proteomic assays are limited due to sample processing biases. Simpler assays can also be informative for go/no-go decisions for larger proteomic studies or to inform a preferred experimental design.

5.4. Study findings

Spheroids were characterized with respect to their relative expression of hepatic marker proteins relative to monolayers as well as across spheroid treatment conditions and time points. The high-level detection of differences in protein expression between groups as shown by differential growth kinetics, progressively different expression profiles as global proteomic adaptations are important considerations within the context of the translational aspects of this study. HepG2 cells, cultured as spheroids, display a wider array of structural and biological variances compared to their monolayer counterparts, including the much longer time for which they can be maintained in culture compared to conventional monolayer cultures These differences are also generally noted across

methods of spheroid culture techniques. Hanging drop models, while beginning with higher seeding densities, reach a growth plateau between 14 and 21 days in culture, a feature not observed in the equivalent culture time frame for micro mould cultures that generate smaller spheroids and are seeded with much lower cell numbers. In general, HepG2 cells did not display evidence of cytotoxicity in response to the induction drug exposure, even at the high 200 μM doses administered to monolayer cultures during dose response curve assays, a concentration at which some level of cytotoxic activity would typically be expected. Similar observations were made in the case of spheroid cultures where long term exposure of a (relatively) high concentration of 5 μM resulted in no significant cytotoxicity profiles, while exposure to far-lower concentrations were able to generate robust and significant changes at the level of the proteome. Additionally, qualitative changes to the proteomes of HepG2 cells between monolayer and spheroid cultures are evident even in relatively insensitive qualitative measurements such as SDS-PAGE.

This study also provides insight into the previously uncharacterised, dynamic nature of 3D culture systems and highlights important considerations essential to the application of biological models in general. Previous reports have indicated that these changes were not robust in spheroids at 10 days in culture [229] . Here it was demonstrated that increased culture time allowed for a continual divergence of the proteome of spheroid cultured cell replicates from starting equivalent monolayer cultures. There is a continuous, reproducible, divergent temporal proteome profile with regards to proteins associated with hepatic phenotype, cell junctions, extracellular matrix, and cell-adhesion molecules. Taken together, these data demonstrate how highly dynamic the proteome of spheroid cultures are and provides a resource for assessing proteomic changes based on cell culture modulation.

Experimentation using quantitative proteomics, was successful in identifying and quantitating in relative terms, over 4800 proteins in biological replicates across the 7 experimental conditions surveyed in this study. The major changes in the spheroid proteome could be coerced further by addition of a low dose drug cocktail, which shows rudimentary evidence of a shift to a phenotype that is less orientated towards proliferation and shows lower priority for the deposition of ECM proteins and potentially some degree of adaptation to metabolic functionality. Perturbations to the proteome differed between the control and induction conditions compared to both the initial monolayers from which the spheroids were grown and to equivalent spheroids from the same paired timepoint. A potential mechanism by which drug exposed spheroid groups further change protein expression profiles is by modulation of HNF4α transcriptional activity as shown by the differential expression of downstream elements of HNF4α activation which were evident at a protein level

This study also demonstrated the presence of drug exposure driven metabolism in the case of CYPA4 with variable activity in CYP's 2D6 and 2C19. However metabolic activity was not detected for CYP's 1A2, 2B6 and 2C9, which may be related to the "high" concentrations of the probe drugs used in the assay protocol.

The adoption of cross-omics assays and bioinformatics analysis methods have proven useful in the elucidation of the underlying mechanisms for the proteomic changes observed and have further served to characterize these changes in an objective and simplified manner. The differential expression of proteins that are involved in metabolic pathways in drug treated spheroid groups, was supported by the increase in drug metabolism seen within these groups. Though these differences are undoubtedly present, without a representative control for *in vivo* counterparts, the value of these changes is difficult to validate, and their relative representativeness is likely still insufficient for accurate recapitulation of hepatocyte xenobiotic metabolism.

5.5. Study limitations

The requirement to adapt ones established strategies in light of new data can become a source of error and ultimately result in the inability to extrapolate one's findings within the context of the larger experimental question. This study was no exception, and the limitations associated with its design will be discussed below.

For the characterisation of hepatic models in the context of PHH it would have been highly beneficial to include a positive control for primary hepatocyte proteomes to serve as a frame of reference for the observed changes within HepG2 cultures as without these, observation can only be made with respect to other HepG2 cultures, thereby limiting the translational power of the study.

For the establishment of true induced spheroid cultures, it may have been more beneficial to include a set of HepG2 spheroid which were cultured in the presence of a prototypical broad spectrum CYP inducer such as carbamazepine or phenytoin rather than the multi drug substrate cocktail, as it still remains uncertain as to whether these cultures would be able to become even more metabolically adapted, perhaps to the extent that CYP enzymes may be quantifiable within proteomic experiments

The use of TMT labelled proteomics, while generally successful within the context of this study incurs great financial costs and time investment. This limited the possibility for further experimentation, or optimisation in terms of protein extraction methods and differing biological or mass spectrometer contexts.

The use of gradient assisted subcellular fractionation may have been useful in the enragement of cellular fractions which are likely to contain CYP enzymes, such as the microsomal fractions.

The use of reverse phase C18 chromatography at all steps may have biased the protein subsets identifiable within this study, leading to selective removal of highly lipophilic peptides that would be representative of membrane binding domains of intrinsic membrane proteins, like the CYP450's.

References

1. Hansen, CW and L Lønstrup, *The rise in life expectancy and economic growth in the 20th century.* The Economic Journal, 2015. **125**(584): p. 838-52.
2. Pisano, DJ and DS Mantus, *FDA regulatory affairs: a guide for prescription drugs, medical devices, and biologics*. 2019: CRC Press.
3. DiMasi, JA, HG Grabowski, and RW Hansen, *The cost of drug development.* New England Journal of Medicine, 2015. **372**.
4. Adams, CP and VV Brantner, *Estimating the cost of new drug development: is it really $802 million?* Health Affairs, 2006. **25**(2): p. 420-28.
5. Khurana, G, A Rohilla, and A Deep, *Drug Development Process and Novel Drugs Approved by FDA for 2017-18.* Applied Clinical Research, Clinical Trials and Regulatory Affairs, 2018. **5**(2): p. 80-98.
6. Terstappen, GC and A Reggiani, *In silico research in drug discovery.* Trends in Pharmacological Sciences, 2001. **22**(1): p. 23-26.
7. Roberts, BR, *Screening informatics: adding value with meta-data structures and visualization tools.* Drug Discovery Today, 2000. **5**: p. 10-14.
8. Rousseaux, CG and WM Bracken, *Overview of Drug Development*, in *Haschek and Rousseaux's Handbook of Toxicologic Pathology*. 2013, Elsevier. p. 647-85.
9. Arrowsmith, J and P Miller, *Trial watch: phase II and phase III attrition rates 2011–2012.* Nature reviews Drug discovery, 2013. **12**(8): p. 569.
10. Hwang, TJ, et al., *Failure of investigational drugs in late-stage clinical development and publication of trial results.* JAMA internal medicine, 2016. **176**(12): p. 1826-33.
11. Crowther, M, *Phase 4 research: what happens when the rubber meets the road?* Hematology 2013, the American Society of Hematology Education Program Book, 2013. **2013**(1): p. 15-18.
12. Fijal, BA, JM Hall, and JS Witte, *Clinical trials in the genomic era: effects of protective genotypes on sample size and duration of trial.* Controlled Clinical Trials, 2000. **21**(1): p. 7-20.
13. Shuster, E, *Fifty years later: the significance of the Nuremberg Code.* New England Journal of Medicine, 1997. **337**(20): p. 1436-40.
14. Mazzoleni, G, D Di Lorenzo, and N Steimberg, *Modelling tissues in 3D: the next future of pharmaco-toxicology and food research?* Genes & nutrition, 2009. **4**(1): p. 13.
15. Kunz-Schughart, LA, et al., *The use of 3-D cultures for high-throughput screening: the multicellular spheroid model.* Journal of biomolecular screening, 2004. **9**(4): p. 273-85.
16. Bissell, MJ, A Rizki, and IS Mian, *Tissue architecture: the ultimate regulator of breast epithelial function.* Current opinion in cell biology, 2003. **15**(6): p. 753.
17. Nath, S and GR Devi, *Three-dimensional culture systems in cancer research: Focus on tumor spheroid model.* Pharmacology & therapeutics, 2016. **163**: p. 94-108.
18. Nunes, AS, et al., *3D tumor spheroids as in vitro models to mimic in vivo human solid tumors resistance to therapeutic drugs.* Biotechnology and bioengineering, 2019. **116**(1): p. 206-26.
19. Lauschke, VM, et al., *Novel 3D culture systems for studies of human liver function and assessments of the hepatotoxicity of drugs and drug candidates.* Chemical Research in Toxicology, 2016. **29**(12): p. 1936-55.

20. Bell, CC, et al., *Comparison of hepatic 2D sandwich cultures and 3D spheroids for long-term toxicity applications: a multicenter study.* Toxicological Sciences, 2018. **162**(2): p. 655-66.
21. Lauschke, VM, et al., *Novel 3D culture systems for studies of human liver function and assessments of the hepatotoxicity of drugs and drug candidates.* Chemical Research in Toxicology, 2016. **29**(12): p. 1936-55.
22. Schmidt, M, et al., *Spheroid-based 3-dimensional culture models: Gene expression and functionality in head and neck cancer.* Oncology reports, 2016. **35**(4): p. 2431-40.
23. O'Brien, LE, MM Zegers, and KE Mostov, *Building epithelial architecture: insights from three-dimensional culture models.* Nature reviews Molecular cell biology, 2002. **3**(7): p. 531-37.
24. Loeb, L, *Ueber die Entstehung von Bindegewebe Leucocyten und roten Blutkorperchen aus Epithel und uber eine Methode, isolierte Gewebsteile zu zuchten*. 1897.
25. Fine, B and G Vunjak-Novakovic, *Shortcomings of animal models and the rise of engineered human cardiac tissue.* ACS Biomaterials Science & Engineering, 2017. **3**(9): p. 1884-97.
26. Massa, A, et al., *Evolution of the experimental models of cholangiocarcinoma.* Cancers, 2020. **12**(8): p. 2308.
27. Lampis, A, et al., *MIR21 drives resistance to heat shock protein 90 inhibition in cholangiocarcinoma.* Gastroenterology, 2018. **154**(4): p. 1066-79. e5.
28. Yan, HH, et al., *A comprehensive human gastric cancer organoid biobank captures tumor subtype heterogeneity and enables therapeutic screening.* Cell stem cell, 2018. **23**(6): p. 882-97. e11.
29. Praharaj, PP, et al., *Circulating tumor cell-derived organoids: Current challenges and promises in medical research and precision medicine.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2018. **1869**(2): p. 117-27.
30. Rae, C, F Amato, and C Braconi, *Patient-Derived Organoids as a Model for Cancer Drug Discovery.* International journal of molecular sciences, 2021. **22**(7): p. 3483.
31. Allen, DD, et al., *Cell lines as in vitro models for drug screening and toxicity studies.* Drug development and industrial pharmacy, 2005. **31**(8): p. 757-68.
32. Gomez-Lechon, M, et al., *Cell lines: a tool for in vitro drug metabolism studies.* Current drug metabolism, 2008. **9**(1): p. 1-11.
33. Deer, EL, et al., *Phenotype and genotype of pancreatic cancer cell lines.* Pancreas, 2010. **39**(4): p. 425-35.
34. Tharmalingam, T, et al., *Characterization of phenotypic and genotypic diversity in subclones derived from a clonal cell line.* Biotechnology Progress, 2018. **34**(3): p. 613-23.
35. Pagano, E, et al., *Preclinical models in oncological pharmacology: Limits and advantages.* 2021.
36. Blau, HM, *Sir John Gurdon: father of nuclear reprogramming.* Differentiation, 2014. **88**(1): p. 10-12.
37. Takahashi, K and S Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* cell, 2006. **126**(4): p. 663-76.
38. Malik, N and MS Rao, *A review of the methods for human iPSC derivation.* Pluripotent stem cells, 2013: p. 23-33.
39. Li, Y, et al., *Engineering-derived approaches for iPSC preparation, expansion, differentiation and applications.* Biofabrication, 2017. **9**(3): p. 032001.
40. Sura, R, J Hutt, and S Morgan, *Opinion on the Use of Animal Models in Nonclinical Safety Assessment: Pros and Cons.* Toxicologic Pathology, 2021: p. 01926233211003498.

41. Blanset, D, J Hutt, and S Morgan, *Current use of animal models of disease for nonclinical safety testing.* Current Opinion in Toxicology, 2020. **23**: p. 11-16.
42. Yoshizato, K and C Tateno, *A mouse with humanized liver as an animal model for predicting drug effects and for studying hepatic viral infection: where to next?* Expert opinion on drug metabolism & toxicology, 2013. **9**(11): p. 1419-35.
43. Borzelleca, JF, *Paracelsus: herald of modern toxicology.* Toxicological Sciences, 2000. **53**(1): p. 2-4.
44. Liebler, DC and FP Guengerich, *Elucidating mechanisms of drug-induced toxicity.* Nature reviews Drug discovery, 2005. **4**(5): p. 410-20.
45. Whitebread, S, et al., *Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development.* Drug discovery today, 2005. **10**(21): p. 1421-33.
46. Lee, WM, *Drug-induced hepatotoxicity.* New England Journal of Medicine, 2003. **349**(5): p. 474-85.
47. Yun, C-H, RA Okerholm, and FP Guengerich, *Oxidation of the antihistaminic drug terfenadine in human liver microsomes. Role of cytochrome P-450 3A (4) in N-dealkylation and C-hydroxylation.* Drug metabolism and disposition, 1993. **21**(3): p. 403-09.
48. Trauner, M, *Meier PJ, Boyer JL.* Molecular pathogenesis of cholestasis. N Engl J Med, 1998. **339**: p. 1217-27.
49. Robin, M-A, et al., *Plasma membrane cytochromes P450 as neoantigens and autoimmune targets in drug-induced hepatitis.* Journal of hepatology, 1997. **26**: p. 23-30.
50. Reed, JC, *Apoptosis-regulating proteins as targets for drug discovery.* Trends in molecular medicine, 2001. **7**(7): p. 314-19.
51. Pessayre, D, et al. *Mitochondria in steatohepatitis*. in *Seminars in liver disease*. 2001. Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ….
52. Jungermann, K and T Keitzmann, *Zonation of parenchymal and nonparenchymal metabolism in liver.* Annual review of nutrition, 1996. **16**(1): p. 179-203.
53. Gebhardt, R, *Metabolic zonation of the liver: regulation and implications for liver function.* Pharmacology & therapeutics, 1992. **53**(3): p. 275-354.
54. Zong, Y and JR Friedman, *Liver development.* Liver Disease in Children, 2014: p. 1-813.
55. Anwar, A, et al., *Role of Mass Media and Public Health Communications in the COVID-19 Pandemic.* Cureus, 2020. **12**(9): p. e10453.
56. Miyajima, A, M Tanaka, and T Itoh, *Stem/progenitor cells in liver development, homeostasis, regeneration, and reprogramming.* Cell stem cell, 2014. **14**(5): p. 561-74.
57. Aizarani, N, et al., *A human liver cell atlas reveals heterogeneity and epithelial progenitors.* Nature, 2019. **572**(7768): p. 199-204.
58. Li, W, L Li, and L Hui, *Cell plasticity in liver regeneration.* Trends in Cell Biology, 2020. **30**(4): p. 329-38.
59. Ackermann, AM, et al., *Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes.* Molecular metabolism, 2016. **5**(3): p. 233-44.
60. Williams, R, *Hepatic metabolism of drugs.* Gut, 1972. **13**(7): p. 579-85.
61. Vavricka, SR, et al., *Interactions of rifamycin SV and rifampicin with organic anion uptake systems of human liver.* Hepatology, 2002. **36**(1): p. 164-72.
62. Gibson, GG and P Skett, *Introduction to drug metabolism*. 2001: Nelson Thornes.

63. Xu, C, CY-T Li, and A-NT Kong, *Induction of phase I, II and III drug metabolism/transport by xenobiotics.* Archives of Pharmacal Research, 2005. **28**(3): p. 249.
64. Shen, G and AN Kong, *Nrf2 plays an important role in coordinated regulation of Phase II drug metabolism enzymes and Phase III drug transporters.* Biopharmaceutics and Drug Disposition, 2009. **30**(7): p. 345-55.
65. Choudhary, D, et al., *Expression patterns of mouse and human CYP orthologs (families 1–4) during development and in different adult tissues.* Archives of Biochemistry and Biophysics, 2005. **436**(1): p. 50-61.
66. Guengerich, FP, *Cytochrome p450 and chemical toxicology.* Chemical Research in Toxicology, 2007. **21**(1): p. 70-83.
67. Rendic, S and FJD Carlo, *Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors.* Drug Metabolism Reviews, 1997. **29**(1-2): p. 413-580.
68. Rendic, S and FJD Carlo, *Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors.* Drug metabolism reviews, 1997. **29**(1-2): p. 413-580.
69. Zanger, UM and M Schwab, *Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation.* Pharmacology and Therapeutics, 2013. **138**(1): p. 103-41.
70. Chai, X, S Zeng, and W Xie, *Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond.* Expert opinion on drug metabolism & toxicology, 2013. **9**(3): p. 253-66.
71. Xia, L, et al., *Cytochrome P450 induction response in tethered spheroids as a three-dimensional human hepatocyte in vitro model.* Journal of Applied Toxicology, 2016. **36**(2): p. 320-29.
72. Sinz, M, G Wallace, and J Sahi, *Current industrial practices in assessing CYP450 enzyme induction: preclinical and clinical.* The AAPS journal, 2008. **10**(2): p. 391-400.
73. Olson, H, et al., *Concordance of the toxicity of pharmaceuticals in humans and in animals.* Regulatory Toxicology and Pharmacology, 2000. **32**(1): p. 56-67.
74. Tuschl, G, B Lauer, and SO Mueller, *Primary hepatocytes as a model to analyze species-specific toxicity and drug metabolism.* Expert opinion on drug metabolism & toxicology, 2008. **4**(7): p. 855-70.
75. Guillouzo, A, *Liver cell models in in vitro toxicology.* Environmental health perspectives, 1998. **106**(suppl 2): p. 511-32.
76. Natale, A, et al., *Technological advancements for the development of stem cell-based models for hepatotoxicity testing.* Archives of toxicology, 2019. **93**(7): p. 1789-805.
77. Gomez-Lechon, M, et al., *In vitro evaluation of potential hepatotoxicity induced by drugs.* Current pharmaceutical design, 2010. **16**(17): p. 1963-77.
78. Gómez-Lechón, MJ, et al., *Long-term expression of differentiated functions in hepatocytes cultured in three-dimensional collagen matrix.* Journal of cellular physiology, 1998. **177**(4): p. 553-62.
79. Sadasivan, SK, et al., *Developing an in vitro screening assay platform for evaluation of antifibrotic drugs using precision-cut liver slices.* Fibrogenesis & tissue repair, 2015. **8**(1): p. 1-9.
80. Sleno, L, E Varesio, and G Hopfgartner, *Determining protein adducts of fipexide: mass spectrometry based assay for confirming the involvement of its reactive metabolite in covalent binding.* Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 2007. **21**(24): p. 4149-57.

81. Guo, L, et al., *Similarities and differences in the expression of drug metabolizing enzymes between human hepatic cell lines and primary human hepatocytes.* Drug Metabolism and Disposition, 2010: p. dmd. 110.035873.
82. Hart, SN, et al., *A comparison of whole genome gene expression profiles of HepaRG cells and HepG2 cells to primary human hepatocytes and human liver tissues.* Drug metabolism and disposition, 2010. **38**(6): p. 988-94.
83. Ramaiahgari, SC, et al., *A 3D in vitro model of differentiated HepG2 cell spheroids with improved liver-like properties for repeated dose high-throughput toxicity studies.* Archives of Toxicology, 2014. **88**(5): p. 1083-95.
84. Pampaloni, F, EG Reynaud, and EH Stelzer, *The third dimension bridges the gap between cell culture and live tissue.* Nature reviews Molecular cell biology, 2007. **8**(10): p. 839-45.
85. Ravi, M, et al., *3D cell culture systems: advantages and applications.* Journal of Cellular Physiology, 2015. **230**(1): p. 16-26.
86. Edmondson, R, et al., *Three-dimensional cell culture systems and their applications in drug discovery and cell-based biosensors.* Assay and drug development technologies, 2014. **12**(4): p. 207-18.
87. Montanez-Sauri, SI, DJ Beebe, and KE Sung, *Microscale screening systems for 3D cellular microenvironments: platforms, advances, and challenges.* Cellular and molecular life sciences, 2015. **72**(2): p. 237-49.
88. Ramamoorthy, P, et al., *Metastatic tumor-in-a-dish, a novel multicellular organoid to study lung colonization and predict therapeutic response.* Cancer research, 2019. **79**(7): p. 1681-95.
89. Ryan, S-L, et al., *Drug discovery approaches utilizing three-dimensional cell culture.* Assay and drug development technologies, 2016. **14**(1): p. 19-28.
90. Dixit, V, et al., *Application of Micropatterned Cocultured Hepatocytes to Evaluate the Inductive Potential and Degradation Rate of Major Xenobiotic Metabolizing Enzymes.* Drug Metabolism and Disposition, 2016. **44**(2): p. 250-61.
91. Schmelzer, E, et al., *Three-dimensional perfusion bioreactor culture supports differentiation of human fetal liver cells.* Tissue Engineering Part A, 2010. **16**(6): p. 2007-16.
92. Bissell, MJ, *The differentiated state of normal and malignant cells or how to define a "normal" cell in culture*, in *International Review of Cytology*. 1981, Elsevier. p. 27-100.
93. Rubin, K, et al., *Substrate adhesion of rat hepatocytes: on the role of fibronectin in cell spreading.* Experimental cell research, 1981. **135**(1): p. 127-35.
94. Sánchez, A, et al., *Fibronectin regulates morphology, cell organization and gene expression of rat fetal hepatocytes in primary culture.* Journal of hepatology, 2000. **32**(2): p. 242-50.
95. Wang, Y, et al., *ECM proteins in a microporous scaffold influence hepatocyte morphology, function, and gene expression.* Scientific reports, 2016. **6**(1): p. 1-13.
96. Mecham, RP, *Overview of extracellular matrix.* Current protocols in cell biology, 2012. **57**(1): p. 10.1. 1-10.1. 16.
97. Rust, R, C Gantner, and ME Schwab, *Pro-and antiangiogenic therapies: current status and clinical implications.* The FASEB Journal, 2019. **33**(1): p. 34-48.
98. Hynes, RO, *The extracellular matrix: not just pretty fibrils.* Science, 2009. **326**(5957): p. 1216-19.
99. Campbell, ID and MJ Humphries, *Integrin structure, activation, and interactions.* Cold Spring Harbor perspectives in biology, 2011. **3**(3): p. a004994.

100. Horton, ER, et al., *Mechanosensitivity of integrin adhesion complexes: role of the consensus adhesome.* Experimental cell research, 2016. **343**(1): p. 7-13.

101. Rozario, T and DW DeSimone, *The extracellular matrix in development and morphogenesis: a dynamic view.* Developmental biology, 2010. **341**(1): p. 126-40.

102. Leivo, I, et al., *Appearance and distribution of collagens and laminin in the early mouse embryo.* Developmental biology, 1980. **76**(1): p. 100-14.

103. Hussey, GS, JL Dziki, and SF Badylak, *Extracellular matrix-based materials for regenerative medicine.* Nature Reviews Materials, 2018. **3**(7): p. 159-73.

104. Hasin, Y, M Seldin, and A Lusis, *Multi-omics approaches to disease.* Genome biology, 2017. **18**(1): p. 1-15.

105. Wilkins, MR, et al., *Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it.* Biotechnology and Genetic Engineering Reviews, 1996. **13**(1): p. 19-50.

106. Domon, B and R Aebersold, *Mass spectrometry and protein analysis.* Science, 2006. **312**(5771): p. 212-17.

107. Aebersold, RH, et al., *Internal amino acid sequence analysis of proteins separated by one-or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose.* Proceedings of the National Academy of Sciences, 1987. **84**(20): p. 6970-74.

108. Licker, V, *Characterisation of human cerebrospinal fluid (CSF) after Tandem Mass Tag (TMT 0) labelling.* 2006.

109. Prince, JT, *Integration and validation of mass spectrometry proteomics data sets.* 2008.

110. Reid, GE and SA McLuckey, *'Top down'protein characterization via tandem mass spectrometry.* Journal of Mass Spectrometry, 2002. **37**(7): p. 663-75.

111. Link, AJ, et al., *Direct analysis of protein complexes using mass spectrometry.* Nature Biotechnology, 1999. **17**(7): p. 676-82.

112. Pappin, DJ, P Hojrup, and AJ Bleasby, *Rapid identification of proteins by peptide-mass fingerprinting.* Current Biology, 1993. **3**(6): p. 327-32.

113. Hochstrasser, DF and J-C Sanchez, *The dynamic range of protein expression: a challenge for proteomic research.* Proteins, 2000. **1109**: p. 4.

114. ThermoFischer, *User Guide: TMT Mass Tagging Kits and Reagents*. 2012. p. 3-4.

115. Aggarwal, K, LH Choe, and KH Lee, *Shotgun proteomics using the iTRAQ isobaric tags.* Briefings in functional genomics & proteomics, 2006. **5**(2): p. 112-20.

116. Trotter, MW, et al., *Improved sub-cellular resolution via simultaneous analysis of organelle proteomics data across varied experimental conditions.* Proteomics, 2010. **10**(23): p. 4213-19.

117. Rauniyar, N and JR Yates III, *Isobaric labeling-based relative quantification in shotgun proteomics.* Journal of Proteome Research, 2014. **13**(12): p. 5293-309.

118. Zhang, F, et al., *Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020.* Proteomics, 2020. **20**(17-18): p. 1900276.

119. Tran, NH, et al., *Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry.* Nature methods, 2019. **16**(1): p. 63-66.

120. Vaudel, M, et al., *Peptide identification quality control.* Proteomics, 2011. **11**(10): p. 2105-14.

121. Huang, DW, BT Sherman, and RA Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nature protocols, 2009. **4**(1): p. 44-57.

122. Tyanova, S and J Cox, *Perseus: a bioinformatics platform for integrative analysis of proteomics data in cancer research*, in *Cancer systems biology*. 2018, Humana Press. p. 133-48.

123. Kristal, BS, et al., *High-performance liquid chromatography separations coupled with coulometric electrode array detectors*, in *Metabolomics*. 2007, Springer. p. 159-74.

124. Zhang, A, et al., *Modern analytical techniques in metabolomics analysis.* Analyst, 2012. **137**(2): p. 293-300.

125. Wishart, DS, *Emerging applications of metabolomics in drug discovery and precision medicine.* Nature reviews Drug discovery, 2016. **15**(7): p. 473-84.

126. Wishart, DS, *Advances in metabolite identification.* Bioanalysis, 2011. **3**(15): p. 1769-82.

127. Duret, C, et al., *Isolation, characterization, and differentiation to hepatocyte-like cells of nonparenchymal epithelial cells from adult human liver.* Stem Cells, 2007. **25**(7): p. 1779-90.

128. Song, MO, J Li, and JH Freedman, *Physiological and toxicological transcriptome changes in HepG2 cells exposed to copper.* Physiological genomics, 2009. **38**(3): p. 386-401.

129. McMillian, MK, et al., *Nile Red binding to HepG2 cells: an improved assay for in vitro studies of hepatosteatosis.* In Vitro & Molecular Toxicology: A Journal of Basic and Applied Research, 2001. **14**(3): p. 177-90.

130. Xue, Y, et al., *Cytotoxicity and apoptosis induced by silver nanoparticles in human liver HepG2 cells in different dispersion media.* Journal of Applied Toxicology, 2016. **36**(3): p. 352-60.

131. Caro, AA and AI Cederbaum, *Oxidative stress, toxicology, and pharmacology of CYP2E1.* Annu. Rev. Pharmacol. Toxicol., 2004. **44**: p. 27-42.

132. Ruskowitz, ER and CA DeForest, *Photoresponsive biomaterials for targeted drug delivery and 4D cell culture.* Nature Reviews Materials, 2018. **3**(2): p. 1-17.

133. Maronpot, RR, et al., *Hepatic enzyme induction: histopathology.* Toxicologic pathology, 2010. **38**(5): p. 776-95.

134. Ennulat, D, et al., *Effects of hepatic drug-metabolizing enzyme induction on clinical pathology parameters in animals and man.* Toxicologic pathology, 2010. **38**(5): p. 810-28.

135. Honkakoski, P, T Sueyoshi, and M Negishi, *Drug-activated nuclear receptors CAR and PXR.* Annals of medicine, 2003. **35**(3): p. 172-82.

136. Wang, H and EL LeCluyse, *Role of orphan nuclear receptors in the regulation of drug-metabolising enzymes.* Clinical pharmacokinetics, 2003. **42**(15): p. 1331-57.

137. Honkakoski, P and M Negishi, *Regulation of cytochrome P450 (CYP) genes by nuclear receptors.* Biochemical Journal, 2000. **347**(2): p. 321-37.

138. Xie, W, et al., *Reciprocal activation of xenobiotic response genes by nuclear receptors SXR/PXR and CAR.* Genes & development, 2000. **14**(23): p. 3014-23.

139. Fritz, A, et al., *Expression of clinically relevant drug-metabolizing enzymes along the human intestine and their correlation to drug transporters and nuclear receptors: An intra-subject analysis.* Basic & clinical pharmacology & toxicology, 2019. **124**(3): p. 245-55.

140. Tassaneeyakul, W, et al., *Caffeine metabolism by human hepatic cytochromes P450: contributions of 1A2, 2E1 and 3A isoforms.* Biochemical Pharmacology, 1994. **47**(10): p. 1767-76.

141. Tassaneeyakul, W, et al., *Specificity of substrate and inhibitor probes for human cytochromes P450 1A1 and 1A2.* Journal of Pharmacology and Experimental Therapeutics, 1993. **265**(1): p. 401-07.

142. Burk, O, et al., *Antimalarial artemisinin drugs induce cytochrome P450 and MDR1 expression by activation of xenosensors pregnane X receptor and constitutive androstane receptor.* Molecular pharmacology, 2005. **67**(6): p. 1954-65.

143. Riganti, C, et al., *Artemisinin induces doxorubicin resistance in human colon cancer cells via calcium-dependent activation of HIF-1α and P-glycoprotein overexpression.* British journal of pharmacology, 2009. **156**(7): p. 1054-66.

144. Zhu, M, et al., *Cytochrome P450 3A-mediated metabolism of buspirone in human liver microsomes.* Drug metabolism and disposition, 2005. **33**(4): p. 500-07.

145. Jägestedt, M and C von Bahr, *Combination of serotonergic agents resulted in severe adverse effects.* Lakartidningen, 2004. **101**(18): p. 1618-19.

146. Zhou, S-F, *Polymorphism of human cytochrome P450 2D6 and its clinical significance.* Clinical pharmacokinetics, 2009. **48**(12): p. 761-804.

147. Yu, A and RL Haining, *Comparative contribution to dextromethorphan metabolism by cytochrome P450 isoforms in vitro: can dextromethorphan be used as a dual probe for both CYP2D6 and CYP3A activities?* Drug metabolism and disposition, 2001. **29**(11): p. 1514-20.

148. Gorski, JC, et al., *Characterization of dextromethorphan N-demethylation by human liver microsomes: contribution of the cytochrome P450 3A (CYP3A) subfamily.* Biochemical pharmacology, 1994. **48**(1): p. 173-82.

149. Von Moltke, LL, et al., *Multiple human cytochromes contribute to biotransformation of dextromethorphan in-vitro: role of CYP2C9, CYP2C19, CYP2D6, and CYP3A.* Journal of pharmacy and pharmacology, 1998. **50**(9): p. 997-1004.

150. Rendic, S, *Summary of information on human CYP enzymes: human P450 metabolism data.* Drug metabolism reviews, 2002. **34**(1-2): p. 83-448.

151. Lutz, JD and N Isoherranen, *Prediction of relative in vivo metabolite exposure from in vitro data using two model drugs: dextromethorphan and omeprazole.* Drug metabolism and disposition, 2012. **40**(1): p. 159-68.

152. Mo, S-L, et al., *New insights into the structural features and functional relevance of human cytochrome P450 2C9. Part I.* Current drug metabolism, 2009. **10**(10): p. 1075-126.

153. Bort, R, et al., *Hepatic metabolism of diclofenac: role of human CYP in the minor oxidative pathways.* Biochemical pharmacology, 1999. **58**(5): p. 787-96.

154. Redfearn, DP, et al., *Esophageal temperature monitoring during radiofrequency ablation of atrial fibrillation.* Journal of cardiovascular electrophysiology, 2005. **16**(6): p. 589-93.

155. Mancy, A, et al., *Diclofenac and its derivatives as tools for studying human cytochromes P450 active sites: particular efficiency and regioselectivity of P450 2Cs.* Biochemistry, 1999. **38**(43): p. 14264-70.

156. Ngui, JS, et al., *Cytochrome P450 3A4-mediated interaction of diclofenac and quinidine.* Drug metabolism and disposition, 2000. **28**(9): p. 1043-50.

157. Klieber, S, et al., *Contribution of the N-glucuronidation pathway to the overall in vitro metabolic clearance of midazolam in humans.* Drug metabolism and disposition, 2008. **36**(5): p. 851-62.

158. Hyland, R, et al., *In vitro and in vivo glucuronidation of midazolam in humans.* British Journal of Clinical Pharmacology, 2009. **67**(4): p. 445-54.

159. Seo, K-A, et al., *Metabolism of 1'-and 4-hydroxymidazolam by glucuronide conjugation is largely mediated by UDP-glucuronosyltransferases 1A4, 2B4, and 2B7.* Drug metabolism and disposition, 2010. **38**(11): p. 2007-13.

160. Liu, Y, et al., *The inhibition study of human UDP-glucuronosyltransferases with cytochrome P450 selective substrates and inhibitors.* Journal of enzyme inhibition and medicinal chemistry, 2011. **26**(3): p. 386-93.

161. Patki, KC, LL Von Moltke, and DJ Greenblatt, *In vitro metabolism of midazolam, triazolam, nifedipine, and testosterone by human liver microsomes and recombinant cytochromes p450: role of cyp3a4 and cyp3a5.* Drug metabolism and disposition, 2003. **31**(7): p. 938-44.

162. Williams, JA, et al., *Comparative metabolic capabilities of CYP3A4, CYP3A5, and CYP3A7.* Drug metabolism and disposition, 2002. **30**(8): p. 883-91.

163. Frick, A, J Kopitz, and N Bergemann, *Omeprazole reduces clozapine plasma concentrations.* Pharmacopsychiatry, 2003. **36**(03): p. 121-23.

164. Krusekopf, S, et al., *Time-dependent transcriptional induction of CYP1A1, CYP1A2 and CYP1B1 mRNAs by H+/K+-ATPase inhibitors and other xenobiotics.* Xenobiotica, 2003. **33**(2): p. 107-18.

165. Harvey, JL, et al., *Effect of the adrenal 11-β-hydroxylase inhibitor metyrapone on human hepatic cytochrome P-450 expression: induction of cytochrome P-450 3A4.* Drug metabolism and disposition, 2000. **28**(1): p. 96-101.

166. Karam, W, et al., *Human CYP2C19 is a major omeprazole 5-hydroxylase, as demonstrated with recombinant cytochrome P450 enzymes.* Drug metabolism and disposition, 1996. **24**(10): p. 1081-87.

167. Furuta, S, et al., *Inhibition of drug metabolism in human liver microsomes by nizatidine, cimetidine and omeprazole.* Xenobiotica, 2001. **31**(1): p. 1-10.

168. Li, X-Q, et al., *Enantiomer/enantiomer interactions between the S-and R-isomers of omeprazole in human cytochrome P450 enzymes: major role of CYP2C19 and CYP3A4.* Journal of Pharmacology and Experimental Therapeutics, 2005. **315**(2): p. 777-87.

169. Foti, RS and JL Wahlstrom, *CYP2C19 inhibition: the impact of substrate probe selection on in vitro inhibition profiles.* Drug metabolism and disposition, 2008. **36**(3): p. 523-28.

170. Zhang, T, Y Zhu, and C Gunaratna, *Rapid and quantitative determination of metabolites from multiple cytochrome P450 probe substrates by gradient liquid chromatography–electrospray ionization-ion trap mass spectrometry.* Journal of Chromatography B, 2002. **780**(2): p. 371-79.

171. Kobayashi, K, et al., *Involvement of CYP2E1 as a low-affinity enzyme in phenacetin O-deethylation in human liver microsomes.* Drug metabolism and disposition, 1999. **27**(8): p. 860-65.

172. Tassaneeyakul, W, et al., *Specificity of substrate and inhibitor probes for human cytochromes P450 1A1 and 1A2.* Journal of Pharmacology and Experimental Therapeutics, 1993. **265**(1): p. 401-07.

173. Fukami, T, et al., *CYP2A13 metabolizes human CYP1A2 substrates, phenacetin and theophilline.* Drug metabolism and disposition, 2006.

174. Rendic, S, *Summary of information on human CYP enzymes: human P450 metabolism data.* Drug metabolism reviews, 2002. **34**(1-2): p. 83-448.

175. Gwon, K, E Kim, and G Tae, *Heparin-hyaluronic acid hydrogel in support of cellular activities of 3D encapsulated adipose derived stem cells.* Acta biomaterialia, 2017. **49**: p. 284-95.

176. Hurrell, T, et al., *Characterization and reproducibility of HepG2 hanging drop spheroids toxicology in vitro.* Toxicology in Vitro, 2018. **50**: p. 86-94.

177. Bosilkovska, M, et al., *Geneva cocktail for cytochrome p450 and P-glycoprotein activity assessment using dried blood spots.* Clinical Pharmacology and Therapeutics, 2014. **96**(3): p. 349-59.

178. Wiechelman, KJ, RD Braun, and JD Fitzpatrick, *Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation.* Analytical biochemistry, 1988. **175**(1): p. 231-37.

179. Vichai, V and K Kirtikara, *Sulforhodamine B colorimetric assay for cytotoxicity screening.* Nature protocols, 2006. **1**(3): p. 1112-16.

180. Croue, A, et al., *Contribution of immunohistochemistry, electron microscopy, and cell culture to the characterization of nonfunctioning pituitary adenomas: a study of 40 cases.* Human pathology, 1992. **23**(12): p. 1332-39.

181. Zhang, H, W Wen, and J Yan, *Application of immunohistochemistry technique in hydrobiological studies.* Aquaculture and fisheries, 2017. **2**(3): p. 140-44.

182. Ellero, AA, et al., *Continual proteomic divergence of HepG2 cells as a consequence of long-term spheroid culture.* Scientific Reports, 2021. **11**(1): p. 1-14.

183. Hamilton, GA, C Westmoreland, and E George, *Effects of medium composition on the morphology and function of rat hepatocytes cultured as spheroids and monolayers.* In Vitro Cellular & Developmental Biology-Animal, 2001. **37**(10): p. 656-67.

184. Krahn, KN, et al., *Fluorescently labeled collagen binding proteins allow specific visualization of collagen in tissues and live cell culture.* Analytical biochemistry, 2006. **350**(2): p. 177-85.

185. Aper, SJ, et al., *Colorful protein-based fluorescent probes for collagen imaging.* PloS One, 2014. **9**(12): p. e114983.

186. Meghezi, S, et al., *Engineering 3D cellularized collagen gels for vascular tissue regeneration.* Journal of visualized experiments: JoVE, 2015(100).

187. Ignatius, A, et al., *Tissue engineering of bone: effects of mechanical strain on osteoblastic cells in type I collagen matrices.* Biomaterials, 2005. **26**(3): p. 311-18.

188. Massimino, LC, et al., *Use of collagen and auricular cartilage in bioengineering: Scaffolds for tissue regeneration.* Cell and Tissue Banking, 2020: p. 1-12.

189. Rajan, N, et al., *Preparation of ready-to-use, storable and reconstituted type I collagen from rat tail tendon for tissue engineering applications.* Nature protocols, 2006. **1**(6): p. 2753-58.

190. Carter, M and JC Shieh, *Guide to research techniques in neuroscience*. 2015: Academic Press.

191. Elaut, G, et al., *Molecular mechanisms underlying the dedifferentiation process of isolated hepatocytes and their cultures.* Current drug metabolism, 2006. **7**(6): p. 629-60.

192. Modriansky, M, et al., *Human hepatocyte-a model for toxicological studies. Functional and biochemical characterization.* General Physiology and Biophysics, 2000. **19**(2): p. 223-36.

193. Chiu, JH, et al., *The formation of bile canaliculi in human hepatoma cell lines.* Hepatology, 1990. **11**(5): p. 834-42.

194. Takezawa, T and M Uzu, *HepG2-NIAS cells, a new subline of HepG2 cells that can enhance not only CYP3A4 activity but also expression of drug transporters and form bile canaliculus-like networks by the oxygenation culture via a collagen vitrigel membrane.* The Journal of Toxicological Sciences, 2022. **47**(1): p. 39-50.

195. Ozeki, M, et al., *Prevention of cell growth by suppression of villin expression in lithocholic acid-stimulated HepG2 cells.* Journal of Histochemistry and Cytochemistry, 2019. **67**(2): p. 129-41.

196. 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. 554-64.
197. Štampar, M, et al., *Characterization of in vitro 3D cell model developed from human hepatocellular carcinoma (HepG2) Cell Line.* Cells, 2020. **9**(12): p. 2557.
198. Kahn, M, *Can we safely target the WNT pathway?* Nature reviews Drug discovery, 2014. **13**(7): p. 513-32.
199. Rowlands, TM, et al., *B-Catenin and cyclin D1: connecting development to breast cancer.* Cell Cycle, 2004. **3**(2): p. 143-46.
200. Mendonsa, AM, T-Y Na, and BM Gumbiner, *E-cadherin in contact inhibition and cancer.* Oncogene, 2018. **37**(35): p. 4769-80.
201. Scientific, TF. *3D cell culture handbook*. 2020; Available from: https://assets.thermofisher.com/TFS-Assets/BID/Handbooks/3d-cell-culture-handbook.pdf.
202. Aviner, R, *The science of puromycin: From studies of ribosome function to applications in biotechnology.* Computational and Structural Biotechnology Journal, 2020. **18**: p. 1074-83.
203. Baues, M, et al., *A collagen-binding protein enables molecular imaging of kidney fibrosis in vivo.* Kidney International, 2020. **97**(3): p. 609-14.
204. Korver, S, et al., *The application of cytokeratin-18 as a biomarker for drug-induced liver injury.* Archives of Toxicology, 2021. **95**(11): p. 3435-48.
205. Schutte, B, et al., *Keratin 8/18 breakdown and reorganization during apoptosis.* Experimental Cell Research, 2004. **297**(1): p. 11-26.
206. Caulín, C, GS Salvesen, and RG Oshima, *Caspase cleavage of keratin 18 and reorganization of intermediate filaments during epithelial cell apoptosis.* The Journal of cell biology, 1997. **138**(6): p. 1379-94.
207. Naci, H, et al., *Raising the bar for market authorisation of new drugs.* Bmj, 2012. **345**: p. e4261.
208. Wieseler, B, N McGauran, and TJB Kaiser, *New drugs: where did we go wrong and what can we do better?* 2019. **366**: p. l4340.
209. Mullard, A, *Parsing clinical success rates.* Nature Reviews Drug Discovery, 2016. **15**(7): p. 447.
210. Harrison, RK, *Phase II and phase III failures: 2013–2015.* Nature Reviews Drug Discovery, 2016. **15**(12): p. 817.
211. Babai, S, L Auclert, and H Le-Louët, *Safety data and withdrawal of hepatotoxic drugs.* Therapie, 2018.
212. Malangu, N, *Pharmacokinetics and Adverse Effects of Drugs: Mechanisms and Risks Factors*. 2018: BoD–Books on Demand.
213. Gómez-Lechón, MJ and L Tolosa, *Human hepatocytes derived from pluripotent stem cells: a promising cell model for drug hepatotoxicity screening.* Archives of toxicology, 2016. **90**(9): p. 2049-61.
214. Wong, MW, et al., *Status and use of Induced Pluripotent Stem Cells (iPSCs) in toxicity testing*, in *Drug-Induced Liver Toxicity*. 2018, Springer. p. 199-212.
215. Heslop, JA, et al., *Mechanistic evaluation of primary human hepatocyte culture using global proteomic analysis reveals a selective dedifferentiation profile.* Archives of toxicology, 2017. **91**(1): p. 439-52.
216. Sampaziotis, F, CP Segeritz, and L Vallier, *Potential of human induced pluripotent stem cells in studies of liver disease.* Hepatology, 2015. **62**(1): p. 303-11.

217. Donato, MT, L Tolosa, and MJ Gómez-Lechón, *Culture and functional characterization of human hepatoma HepG2 cells*, in *Protocols in In Vitro Hepatocyte Research*. 2015, Springer. p. 77-93.

218. Baxter, M, et al., *Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes.* Journal of Hepatology, 2015. **62**(3): p. 581-89.

219. Zabulica, M, et al., *Guide to the Assessment of Mature Liver Gene Expression in Stem Cell-Derived Hepatocytes.* Stem cells and development, 2019(ja).

220. Fennema, E, et al., *Spheroid culture as a tool for creating 3D complex tissues.* Trends in biotechnology, 2013. **31**(2): p. 108-15.

221. Huang, B-W and J-Q Gao, *Application of 3D cultured multicellular spheroid tumor models in tumor-targeted drug delivery system research.* Journal of controlled release, 2018. **270**: p. 246-59.

222. Duval, K, et al., *Modeling physiological events in 2D vs. 3D cell culture.* Physiology, 2017. **32**(4): p. 266-77.

223. Weyers, C, et al., *Use of a non-hepatic cell line highlights limitations associated with cell-based assessment of metabolically induced toxicity.* Drug and chemical toxicology, 2019: p. 1-7.

224. Shah, U-K, et al., *A three-dimensional in vitro HepG2 cells liver spheroid model for genotoxicity studies.* Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 2018. **825**: p. 51-58.

225. Senthilraja, P and K Kathiresan, *In vitro cytotoxicity MTT assay in Vero, HepG2 and MCF-7 cell lines study of Marine Yeast.* J Appl Pharm Sci, 2015. **5**(3): p. 080-84.

226. Choi, JM, et al., *HepG2 cells as an in vitro model for evaluation of cytochrome P450 induction by xenobiotics.* Archives of pharmacal research, 2015. **38**(5): p. 691-704.

227. Ramaiahgari, SC, et al., *From the cover: Three-dimensional (3D) HepaRG spheroid model with physiologically relevant xenobiotic metabolism competence and hepatocyte functionality for liver toxicity screening.* Toxicological Sciences, 2017. **159**(1): p. 124-36.

228. Reynolds, P, *In sickness and in health: the importance of translational regulation.* Archives of disease in childhood, 2002. **86**(5): p. 322-24.

229. Hurrell, T, KS Lilley, and AD Cromarty, *Proteomic responses of HepG2 cell monolayers and 3D spheroids to selected hepatotoxins.* Toxicology Letters, 2019. **300**: p. 40-50.

230. Asthana, A and WS Kisaalita, *Is time an extra dimension in 3D cell culture?* Drug discovery today, 2016. **21**(3): p. 395-99.

231. Tyanova, S, et al., *The Perseus computational platform for comprehensive analysis of (prote) omics data.* Nature methods, 2016. **13**(9): p. 731-40.

232. Polpitiya, AD, et al., *DAnTE: a statistical tool for quantitative analysis of-omics data.* Bioinformatics, 2008. **24**(13): p. 1556-58.

233. Team, R, *RStudio: integrated development for R.* RStudio, Inc., Boston, MA URL http://www. rstudio. com, 2015. **42**(14): p. 84.

234. Wickham, H, *ggplot2: Elegant Graphics for Data Analysis*. 2016: Springer-Verlag New York.

235. Parviz, F, et al., *Hepatocyte nuclear factor 4α controls the development of a hepatic epithelium and liver morphogenesis.* Nature genetics, 2003. **34**(3): p. 292-96.

236. Kyrmizi, I, et al., *Plasticity and expanding complexity of the hepatic transcription factor network during liver development.* Genes & development, 2006. **20**(16): p. 2293-305.

237. Morimoto, A, et al., *An HNF4α–microRNA-194/192 signaling axis maintains hepatic cell function.* Journal of Biological Chemistry, 2017. **292**(25): p. 10574-85.

238. Wang, X and Q Wang, *Alpha-fetoprotein and hepatocellular carcinoma immunity.* Canadian Journal of Gastroenterology and Hepatology, 2018. **2018**.

239. Sala-Trepat, JM, et al., *Changes in expression of albumin and. alpha.-fetoprotein genes during rat liver development and neoplasia.* Biochemistry, 1979. **18**(11): p. 2167-78.

240. Galle, PR, et al., *Biology and significance of alpha-fetoprotein in hepatocellular carcinoma.* Liver International, 2019. **39**(12): p. 2214-29.

241. Fuchs, S, A Hofmann, and CJ Kirkpatrick, *Microvessel-like structures from outgrowth endothelial cells from human peripheral blood in 2-dimensional and 3-dimensional co-cultures with osteoblastic lineage cells.* Tissue Engineering, 2007. **13**(10): p. 2577-88.

242. Antoni, D, et al., *Three-dimensional cell culture: a breakthrough in vivo.* International Journal of Molecular Sciences, 2015. **16**(3): p. 5517-27.

243. Van Duinen, V, et al., *Microfluidic 3D cell culture: from tools to tissue models.* Current Opinion in Biotechnology, 2015. **35**: p. 118-26.

244. Schwartz, R, et al., *Pluripotent stem cell-derived hepatocyte-like cells.* Biotechnology Advances, 2014. **32**(2): p. 504-13.

245. Sun, Y, et al., *Loss of alanine-glyoxylate and serine-pyruvate aminotransferase expression accelerated the progression of hepatocellular carcinoma and predicted poor prognosis.* Journal of translational medicine, 2019. **17**(1): p. 1-16.

246. Mischoulon, D, et al., *Growth-dependent inhibition of CCAAT enhancer-binding protein (C/EBP alpha) gene expression during hepatocyte proliferation in the regenerating liver and in culture.* Molecular and cellular biology, 1992. **12**(6): p. 2553-60.

247. Kim, S-H, J Turnbull, and S Guimond, *Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor.* The Journal of endocrinology, 2011. **209**(2): p. 139-51.

248. Mi, H, et al., *Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v. 14.0).* Nature protocols, 2019. **14**(3): p. 703.

249. Hsu, DK, et al., *Galectin-3 expression is induced in cirrhotic liver and hepatocellular carcinoma.* International journal of cancer, 1999. **81**(4): p. 519-26.

250. Margadant, C, et al., *Epigenetic regulation of galectin-3 expression by β1 integrins promotes cell adhesion and migration.* Journal of Biological Chemistry, 2012. **287**(53): p. 44684-93.

251. Casagrande, V, et al., *Hepatocyte specific TIMP3 expression prevents diet dependent fatty liver disease and hepatocellular carcinoma.* Scientific reports, 2017. **7**(1): p. 1-12.

252. Ong, J, et al., *Imaging-based screen identifies laminin 411 as a physiologically relevant niche factor with importance for i-Hep applications.* Stem cell reports, 2018. **10**(3): p. 693-702.

253. Garcia, MA, WJ Nelson, and N Chavez, *Cell–Cell Junctions Organize Structural and Signaling Networks.* Cold Spring Harbor perspectives in biology, 2018. **10**(4): p. a029181.

254. Sakai, Y, S Yamagami, and K Nakazawa, *Comparative analysis of gene expression in rat liver tissue and monolayer-and spheroid-cultured hepatocytes.* Cells Tissues Organs, 2010. **191**(4): p. 281-88.

255. Fennema, E, et al., *Spheroid culture as a tool for creating 3D complex tissues.* 2013. **31**(2): p. 108-15.

256. Bierwolf, J, et al., *Primary rat hepatocyte culture on 3D nanofibrous polymer scaffolds for toxicology and pharmaceutical research.* 2011. **108**(1): p. 141-50.

257. Kim, Y and PJPo Rajagopalan, *3D hepatic cultures simultaneously maintain primary hepatocyte and liver sinusoidal endothelial cell phenotypes.* 2010. **5**(11): p. e15456.

258. Winter, A, et al., *Cartilage-like gene expression in differentiated human stem cell spheroids: A comparison of bone marrow–derived and adipose tissue–derived stromal cells.* 2003. **48**(2): p. 418-29.

259. Turner, PA, et al., *Spheroid Culture System Confers Differentiated Transcriptome Profile and Functional Advantage to 3T3-L1 Adipocytes.* 2018. **46**(5): p. 772-87.

260. Karsdal, MA, et al., *Novel insights into the function and dynamics of extracellular matrix in liver fibrosis.* American Journal of Physiology-Gastrointestinal and Liver Physiology, 2015. **308**(10): p. G807-G30.

261. Davidenko, N, et al., *Evaluation of cell binding to collagen and gelatin: a study of the effect of 2D and 3D architecture and surface chemistry.* 2016. **27**(10): p. 148.

262. Ramaiahgari, SC, et al., *A 3D in vitro model of differentiated HepG2 cell spheroids with improved liver-like properties for repeated dose high-throughput toxicity studies.* 2014. **88**(5): p. 1083-95.

263. Lauschke, VM, et al., *Novel 3D culture systems for studies of human liver function and assessments of the hepatotoxicity of drugs and drug candidates.* Chemical Research in Toxicology, 2016. **29**(12): p. 1936-55.

264. Bell, CC, et al., *Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease.* 2016. **6**: p. 25187.

265. Christoforou, A, et al., *A draft map of the mouse pluripotent stem cell spatial proteome.* Nature communications, 2016. **7**(1): p. 1-12.

266. Searle, BC, *Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies.* Proteomics, 2010. **10**(6): p. 1265-69.

267. Kuwajima, T, et al., *ClearT: a detergent-and solvent-free clearing method for neuronal and non-neuronal tissue.* 2013. **140**(6): p. 1364-68.

268. Berger, B, et al., *Comparison of liver cell models using the basel phenotyping cocktail.* Frontiers in Pharmacology, 2016. **7**: p. 443.

269. Resing, KA, et al., *Improving reproducibility and sensitivity in identifying human proteins by shotgun proteomics.* Analytical Chemistry, 2004. **76**(13): p. 3556-68.

270. Luckert, C, et al., *Comparative analysis of 3D culture methods on human HepG2 cells.* Archives of Toxicology, 2017. **91**(1): p. 393-406.

271. Wrzesinski, K, et al., *HepG2/C3A 3D spheroids exhibit stable physiological functionality for at least 24 days after recovering from trypsinisation.* 2013. **2**(3): p. 163-72.

272. Fey, SJ and K Wrzesinski, *Determination of drug toxicity using 3D spheroids constructed from an immortal human hepatocyte cell line.* Toxicological Sciences, 2012. **127**(2): p. 403-11.

273. Aweeka, FT and PI German, *Clinical pharmacology of artemisinin-based combination therapies.* Clinical Pharmacokinetics, 2008. **47**(2): p. 91-102.

274. Wong, J and K Yuen, *Improved oral bioavailability of artemisinin through inclusion complexation with β-and γ-cyclodextrins.* International Journal of Pharmaceutics, 2001. **227**(1-2): p. 177-85.

275. Mahmood, I and C Sahajwalla, *Clinical pharmacokinetics and pharmacodynamics of buspirone, an anxiolytic drug.* Clinical Pharmacokinetics, 1999. **36**(4): p. 277-87.

276. Loane, C and M Politis, *Buspirone: what is it all about?* Brain Research, 2012. **1461**: p. 111-18.

277. Pope, LE, et al., *Pharmacokinetics of dextromethorphan after single or multiple dosing in combination with quinidine in extensive and poor metabolizers.* The Journal of Clinical Pharmacology, 2004. **44**(10): p. 1132-42.
278. Kazis, A, V Kimiskidis, and I Niopas, *Pharmacokinetics of dextromethorphan and dextrorphan in epileptic patients.* Acta Neurologica Scandinavica, 1996. **93**(2-3): p. 94-98.
279. Lissy, M, et al., *Pharmacokinetic comparison of an oral diclofenac potassium liquid-filled soft gelatin capsule with a diclofenac potassium tablet.* Expert Opinion on Pharmacotherapy, 2010. **11**(5): p. 701-08.
280. Knoester, P, et al., *Pharmacokinetics and pharmacodynamics of midazolam administered as a concentrated intranasal spray. A study in healthy volunteers.* British Journal of Clinical Pharmacology, 2002. **53**(5): p. 501-07.
281. Haschke, M, et al., *Pharmacokinetics and pharmacodynamics of nasally delivered midazolam.* British Journal of Clinical Pharmacology, 2010. **69**(6): p. 607-16.
282. Mostafavi, SA and N Tavakoli, *Relative bioavailability of omeprazole capsules after oral dosing.* DARU Journal of Pharmaceutical Sciences, 2004. **12**(4): p. 146-50.
283. Raaflaub, J and U Dubach, *On the pharmacokinetics of phenacetin in man.* European journal of clinical pharmacology, 1975. **8**(3): p. 261-65.
284. Allaire, J, *RStudio: integrated development environment for R.* Boston, MA, 2012. **770**(394): p. 165-71.
285. Team, RC, *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* http://www.R-project.org/,2013.
286. Love, MI, W Huber, and S Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biology, 2014. **15**(12): p. 1-21.
287. Ashburner, M, et al., *Gene ontology: tool for the unification of biology.* Nature Genetics, 2000. **25**(1): p. 25-29.
288. Liberzon, A, et al., *Molecular signatures database (MSigDB) 3.0.* Bioinformatics, 2011. **27**(12): p. 1739-40.
289. Gillespie, M, et al., *The reactome pathway knowledgebase 2022.* Nucleic Acids Research, 2022. **50**(D1): p. D687-D92.
290. Kanehisa, M and S Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Research, 2000. **28**(1): p. 27-30.
291. Wickham, H, et al., *Welcome to the Tidyverse.* Journal of open source software, 2019. **4**(43): p. 1686.
292. Wickham, H, et al. *dplyr*. in *useR! Conference*. 2014.
293. Wickham, H, *Data analysis*, in *ggplot2*. 2016, Springer. p. 189-201.
294. Korotkevich, G, et al., *Fast gene set enrichment analysis.* BioRxiv, 2021: p. 060012.
295. Hänzelmann, S, R Castelo, and J Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data.* BMC Bioinformatics, 2013. **14**(1): p. 1-15.
296. Luo, W and C Brouwer, *Pathview: an R/Bioconductor package for pathway-based data integration and visualization.* Bioinformatics, 2013. **29**(14): p. 1830-31.
297. Luo, W, et al., *GAGE: generally applicable gene set enrichment for pathway analysis.* BMC Bioinformatics, 2009. **10**(1): p. 1-17.

298. Walter, W, F Sánchez-Cabo, and M Ricote, *GOplot: an R package for visually combining expression data with functional analysis.* Bioinformatics, 2015. **31**(17): p. 2912-14.

299. Kolde, R, *Pheatmap: pretty heatmaps.* R package version, 2012. **1**(2): p. 726.

300. Blighe, K, S Rana, and M Lewis, *EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.* R package version, 2019. **1**(0).

301. Zhu, X, et al., *Ultrafast optical clearing method for three-dimensional imaging with cellular resolution.* Proceedings of the National Academy of Sciences, 2019. **116**(23): p. 11480-89.

302. Leuschner, M, *ASSESSING INTER-METHOD AGREEMENT OF DRUG-BASED PHENOTYPING METRICS BETWEEN DRIED BLOOD SPOT AND PLASMA SAMPLING*, in *Pharmacology*. 2019, University of Pretoria: Online Repository.

303. Spaggiari, D, et al., *A cocktail approach for assessing the in vitro activity of human cytochrome P450s: an overview of current methodologies.* Journal of Pharmaceutical and Biomedical Analysis, 2014. **101**: p. 221-37.

304. Wrzesinski, K, et al., *HepG2/C3A 3D spheroids exhibit stable physiological functionality for at least 24 days after recovering from trypsinisation.* Toxicology Research, 2013. **2**(3): p. 163-72.

305. Schubert, OT, et al., *Quantitative proteomics: challenges and opportunities in basic and applied research.* Nature Protocols, 2017. **12**(7): p. 1289-94.

306. Pursiheimo, A, et al., *Optimization of statistical methods impact on quantitative proteomics data.* Journal of Proteome Research, 2015. **14**(10): p. 4118-26.

307. Zhu, Y, et al., *DEqMS: a method for accurate variance estimation in differential protein expression analysis.* Molecular & Cellular Proteomics, 2020. **19**(6): p. 1047-57.

308. Lazar, C, *ImputeLCMD: A Collection of Methods for Left-Censored Missing Data Imputation, 2015.* Google Scholar There is no corresponding record for this reference.

309. Lin, M-H, et al., *Benchmarking differential expression, imputation and quantification methods for proteomics data.* Briefings in Bioinformatics, 2022. **23**(3): p. bbac138.

310. Hiort, P, et al., *multiFLEX-LF: A Computational Approach to Quantify the Modification Stoichiometries in Label-Free Proteomics Data Sets.* Journal of proteome research, 2022. **21**(4): p. 899-909.

311. Langley, SR and M Mayr, *Comparative analysis of statistical methods used for detecting differential expression in label-free mass spectrometry proteomics.* Journal of Proteomics, 2015. **129**: p. 83-92.

312. Ferreira, JA, *The Benjamini-Hochberg method in the case of discrete test statistics.* The international journal of biostatistics, 2007. **3**(1).

313. Benjamini, Y and Y Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal statistical society: series B (Methodological), 1995. **57**(1): p. 289-300.

314. Thomas, PD, et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Research, 2003. **13**(9): p. 2129-41.

315. Pomaznoy, M, B Ha, and B Peters, *GOnet: a tool for interactive Gene Ontology analysis.* BMC Bioinformatics, 2018. **19**(1): p. 1-8.

316. Bauer, S, et al., *Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration.* Bioinformatics, 2008. **24**(14): p. 1650-51.

317. Al-Shahrour, F, R Díaz-Uriarte, and J Dopazo, *FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.* Bioinformatics, 2004. **20**(4): p. 578-80.

318. Eden, E and R Navon, *ISDL & Yakhini, Z.(2009). Gorilla: A tool for discovery and visualization of enriched go terms in ranked gene lists.* BMC bioinformatics. **10**: p. 48.
319. Subramanian, A, et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545-50.
320. Reimand, J, et al., *Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap.* Nature Protocols, 2019. **14**(2): p. 482-517.
321. Lavallée-Adam, M, et al., *PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data.* Journal of Proteome Research, 2014. **13**(12): p. 5496-509.
322. Sumi, K, et al., *Cooperative interaction between hepatocyte nuclear factor 4α and GATA transcription factors regulates ATP-binding cassette sterol transporters ABCG5 and ABCG8.* Molecular and Cellular Biology, 2007. **27**(12): p. 4248-60.
323. Qi, T, T Dunlap, and Y Cao, *Embracing Project Optimus: Can we Leverage Evolutionary Theory to Optimize Dosing in Oncology?* Pharmaceutical Research, 2022: p. 1-7.
324. Prakash, A, et al., *Assessing bias in experiment design for large scale mass spectrometry-based quantitative proteomics.* Molecular & Cellular Proteomics, 2007. **6**(10): p. 1741-48.
325. Achour, B, et al., *Global proteomic analysis of human liver microsomes: rapid characterization and quantification of hepatic drug-metabolizing enzymes.* Drug Metabolism and Disposition, 2017. **45**(6): p. 666-75.