Proteomic assessment of potential *in vitro* hepatotoxicity models

by

Tracey Hurrell

A thesis submitted in fulfilment of the requirements for the degree

Doctorate of Philosophy

in

Pharmacology

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

Supervisor: Professor AD Cromarty (University of Pretoria)

Co-supervisor: Professor KS Lilley (University of Cambridge)

2016

Declaration

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

I,	Tracey Hurrell
Student number	27177476
Subject of work	Thesis
Title	"Proteomic assessment of potential *in vitro* hepatotoxicity models"

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 applicable 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 for their contribution and support:

- My supervisor, Professor Duncan Cromarty, for his guidance, support, advice, understanding and belief in my ability
- The National Research Foundation and the Thuthuka Funding Instrument (PhD Track 2013 - 2016, TTK13061819300) without which this project would not have been financially feasible (opinions expressed within are those of the author and can not to be attributed to the NRF)
- Professor Kathryn Lilley who supported both failed and successful Commonwealth applications and created countless opportunities for me to excel
- The Commonwealth Scholarship Commission in the UK for awarding a Commonwealth Scholarship to attend the University of Cambridge as a split-site PhD scholar (ZACS-2014-653)
- The Commonwealth European and International Cambridge Trust who contributed to my University fees at the University of Cambridge
- The Cambridge Centre for Proteomics and the Department of Biochemistry, University of Cambridge for hosting and teaching me invaluable skills throughout my scholarship
- Dr Claire Mulvey and Dr Andy Christoforou who gave of their time to encourage and train me in all things proteomics
- Professor Ludovic Vallier and the Stem Cell Institute at the Anne McLaren laboratory, University of Cambridge who selflessly hosted me as a visitor in his laboratory
- Dr Charis-Patricia Segeritz who generously gave of her time to enlighten and educate me on the wonderful world of stem cell research
- Lucy Cavendish College and all of its members for being a true support system and providing opportunities for personal growth and enrichment
- Hughes Hall Lucy Cavendish College Boat Club and TUKS Rowing Club who taught me how to row, be a member of a team, overcome limitations and strive to be better
- Personnel, colleagues and friends from the Department of Pharmacology for their support
- My friends who have provided comical relief and encouragement
- My parents, sister, brother and nephew. Their love, faith and sacrifice have been indispensable in my successes

"Research is to see what everybody else has seen, and to think what nobody else has thought."

- Albert Szent-Györgyi (1893 - 1986)

Abstract

Purpose

Scientifically credible and valid biological systems are essential in pharmaceutical research and development. Standardizing *in vitro* preclinical hepatotoxicity is confounded by the diversity of origin of cells and the ability to retain hepatocellular function. Key determinants of valid hepatotoxicity models are the resemblance to primary human hepatocytes (PHHs), adaptability to high-throughput screening and biological applicability. Numerous *in vitro* models, including immortalized cell lines and hepatocyte-like cells (HLCs) derived from induced pluripotent stem cells (iPSCs), attempt to reflect features of PHHs. Additional influencing factors are the mechanical and geometric environment which dictate functionality and suggest a role for spatial organisation as a requirement for mimicking PHHs. Due to poor correlation between the cellular genome and proteome, assessing the hepatic phenotypes using proteomics is essential to capture functional cellular responses. The aim of this research was to determine proteomic differences between PHHs and differentially cultured and sourced human hepatocyte-derived cell lines or differentiated HLCs. Additionally, two cellular hepatocyte models were used to generate non-specific, proteome-wide information associated with exposure to selected known hepatotoxins to identify potential proteomic signatures of hepatotoxicity.

Methods

Human iPSCs were generated from fibroblasts and cultured. Stable iPSC colonies were differentiated using a stepwise protocol, over a period of 35 days. Differentiation commenced with endoderm commitment followed by anterior definitive endoderm specification, hepatocyte commitment and hepatocyte maturation to HLCs. Whole cell protein lysates were collected throughout the time-course of differentiation and quantitative proteomics was used to track hepatic markers throughout. Fifty micrograms of sample protein was reduced, alkylated, precipitated, digested with trypsin and labelled using 6-plex tandem-mass tags (TMT). Samples were analysed using nano-LC/MS/MS. Peak lists were searched against a UniProtKB/Swiss-Prot human database using SearchGUI version 2.3.1 with X!Tandem, MS-GF+ and Comet search engines.

Cultured human hepatocellular carcinoma cells (HepG2) were seeded as hanging drops to produce a single, uniform, size controlled spheroid per well (10 000 cells/spheroid; 45 µl). Spheroids were characterised using morphology, protein content, flow cytometry and fluorescence microscopy prior to proteomics. Differentiated HLCs, HepG2 cell monolayers and spheroids as well as commercially available pooled PHHs were cultured and collected as whole cell protein lysates. Replicates were labelled using 6-plex TMT and analysed to determine the proteomic equivalence to PHHs.

Lastly, HepG2 cell monolayers and spheroids were cultured for four days prior to exposure to sub-toxic concentrations of known hepatotoxic drugs Bromfenac, Ketoconazole and Troglitazone for two cycles of 72 hours. A non-hepatotoxic drug, Diphenhydramine, was included for the purposes of validation. Replicates were labelled using 6-plex TMT to determine the proteomic profiles of each drug exposure.

Results

HLC maturity exists on a scale from undifferentiated foetal hepatocytes to differentiated primary adult hepatocytes. Inductive cues in differentiation promote sequential acquisition of hepatocyte specific markers. These protein markers, along with others, were identified throughout the differentiation process. Acquisition of a hepatic phenotype was demonstrated via proteomic characterisation with hepatocyte specific proteins generally peaking between Day 30 and 32 of differentiation.

Comparing cells using a quantitative proteomics approach involving differential stable isotope labelling, identified and quantified over 6 000 proteins per replicate with an overlap of 5 231 proteins. This provided biological insight into the feasibility of using HLCs, HepG2 monolayers and spheroids as a replacement for PHH *in vitro*. The degree to which HLCs, HepG2 monolayers and spheroids mimicked PHH was demonstrated using hierarchical clustering and principal component analysis. Of the proteins identified 1 637, 1 809 and 862 were differentially regulated in HLCs, HepG2 monolayers and spheroids respectively compared to PHH. The HepG2 cell proteome was altered after 10 days of spheroid culture with approximately 34% of proteins differentially regulated across 6 biological replicates. However, following exposure to sub-toxic concentrations of drugs, proteomic profiles were minimally altered. This suggests that changes in the proteome upon spheroid formation may be related to features other than the metabolic components required to identify hepatotoxicity.

Conclusion

The need to replace PHHs in pharmaceutical applications emphasizes the necessity to generate mature hepatic phenotypes as alternatives. HLCs, while currently more representative of a foetal phenotype, have great potential if a mature liver proteome is established. HLCs correlated more closely to PHH than the other models despite some shortcomings in liver specific functions. HepG2 cells appear to be the most limited hepatic phenotype. However, a phenotype closer to PHHs was obtained when cultured as spheroids. Three-dimensional culture is increasing in popularity with hanging drop technology being compatible with many downstream applications. Despite the ability to alter the proteome based on culture technique, proteomic profiling was unable to elucidate drug-specific profiles of hepatotoxicity. Therefore, the complexity of HepG2 spheroids appears insufficient to provide insight into drug-induced hepatotoxicity of the drugs studied here.

Table of Contents

Abbreviations, Acronyms and Symbols
(Limited inclusion of protein acronyms)

A

A1AT	Alpha-1-antitrypsin
Ac-DEVD-AMC	Caspase-3/7 substrate
AcN	Acetonitrile
ASC	Adult stem cells
ADE	Anterior definitive endoderm
ADR	Adverse drug reaction
ADH	Alcohol dehydrogenase
AFP	Alpha-fetoprotein
AGC	Automatic gain control
ALB	Albumin
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
APOF	Apolipoprotein F
APOP	Apoptosis
ANOVA	Analysis of variance
ASGPR	Asialoglycoprotein receptor
AST	Aspartate aminotransferase
ATCC	American Tissue Culture Collection
ATP	Adenosine triphosphate

B

BCA	Bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
BMP	Bone morphogenetic protein
BSEP	Bile salt export pump

C

C18	Silica-based 18 carbon length chain bonded phase
CCP	Cambridge Centre for Proteomics
CDM	Chemically-defined medium
CDM-PVA	Chemically-defined medium with polyvinyl alcohol
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CK18	Cytokeratin 18
CID	Collision-induced dissociation
CO_2	Carbon dioxide
Ct	Threshold cycle
CV	Cell viability / cell enumeration
CXCR	CXC chemokine receptor
CYP	Cytochrome
CYP450	Cytochrome P450

D

dd	Data dependent
ddH_20	Double distilled water
dECM	Decellularized extracellular matrix
DMEM	Dulbecco's Modified Eagle Medium
dH_2O	Distilled water
DILI	Drug-induced liver injury
dNTP	Deoxynucleotide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dsDNA	Double-stranded DNA
DTT	Dithiothreitol

E

ECACC	European Collection of Authenticated Cell Cultures
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGTA	Ethyleneglycoltetraacetic acid
EMEM	Eagle's Minimum Essential Medium
ESC	Embryonic stem cell
ESI	Electrospray ionization

F

FAA	Fatty acid accumulation
FBS	Foetal bovine serum
FDA	Federal Drug Administration
FDR	False discovery rate
FGF	Fibroblast growth factor
FOX	Forkhead box protein

G

GATA	GATA (coding region) binding protein
GDH	Glutamate dehydrogenase
GGT	Gamma-glutamyltransferase
GSH	Reduced glutathione
GST	Glutathione-S-transferase

H

H_2-DCF-DA	Reactive oxygen species substrate
HBSS	Hank's balances salt solution
HCD	Higher energy collisional dissociation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HepG2	Human hepatoma cells

HEX	Homeobox protein
HGF	Hepatocyte growth factor
HLC	Hepatocyte-like cell
HLCA	Hepatocyte comparisons
HLCLTC	Hepatocyte-like cell maturation time course (Day 16 - Day 40)
HLCTC	Hepatocyte-like cell time course (Day 1 - Day 35)
HMGB	High mobility group box 1
hiPSC	Human induced pluripotent stem cell
HNF	Hepatocyte nuclear factor
HPLC	High-performance liquid chromatography
HR/AM-SIM	High-resolution, accurate mass selected ion monitoring scans
HSA	Human serum albumin
HSP	Heat shock protein
HTS	High-throughput screening

I

IAA	Iodoacetamide
ICAT	Isotope-coded affinity tag
IGD	In-gel digestion
ICH	International Conference on Harmonisation
IMDM	Iscove's Modified Dulbecco's Media
IMDs	Inherent metabolic disorders
INDs	Investigational new drugs
IPA	Luciferin-isopropyl acetal
iPSC	Induced pluripotent stem cell
IS	Internal standard
IT	Injection time
iTRAQ	Isobaric tags for relative and absolute quantitation

J

JC-1	Mitochondrial membrane potential substrate

K

KLF-4	Kruppel-like factor-4
KOSR	Knockout serum replacement

L

LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LRH	Liver receptor homolog
LRM	Laboratory for Regenerative Medicine

M

MALDI-TOF	Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry
MAPK	Mitogen-activated protein kinases
MEF	Mouse embryonic fibroblast
MEM-NEAA	Minimum essential medium non-essential amino acids
MET	Mesenchymal-epithelial transition
MGF	Mascot generic format
miR	micro-RNA
MOI	Multiplicity of infection
MMP	Mitochondrial membrane potential
MPT	Mitochondrial permeability transition
mRNA	Messenger ribonucleic acid
MRP	Multidrug resistance-associated protein
MS	Mass spectra, mass spectrometric, mass spectrometry
MS1	Precursor ion scans
MS2	Product ions scans
MuDPIT	Multidimensional protein identification technology

N

NC	Negative control
NCE	Normalized collision energy
NCEs	New chemical entities
NC pump	Binary high-pressure gradient pump
NCBI	National Centre for Biotechnology Information
NEAA	Non-essential amino acids
NHS	N-hydroxy-succinimide
NSAID	Non-steroidal anti-inflammatory drug
5'-NT	5-Nucleotidase

O

OATs	Organic anion transporters
OCT-4	Octamer-binding transcription factor 4
OSM	Oncostatin M

P

PBS	Phosphate buffered saline
PC	Positive control
PDA	Photodiode array
pH	Power of hydrogen (negative logarithmic value of hydrogen ion concentration)
PHH	Primary human hepatocyte
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post-translational modification
PVA	Polyvinyl alcohol

Q

QC	Quality control
QE	Q Exactive Hybrid Quadrupole Orbitrap Mass Spectrometer
qPCR	Quantitative polymerase chain reaction

R

R&D	Research and development
RAFT	Real architecture for 3D tissues
rBM	Reconstituted basement membrane
RCCS	Rotary cell culture system
RGD	Arginylglycylaspartic acid
RIPA	Radioimmunoprecipitation assay 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
RPMI	Roswell Park Memorial Institute Medium
RSA	Republic of South Africa
RSD	Relative standard deviation
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
RT	Reverse transcription
RWV	Rotating wall vessels

S

SCX	Strong cation exchange
SDH	Sorbitol dehydrogenase
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
SIM	Selected ion monitoring
SNCE	Stepped normalized collision energy
SRB	Sulforhodamine B
SOX	SRY (sex determining region Y)-box
ssDNA	Single-stranded DNA
SSEA	Stage-specific embryonic antigen
STs	Sulfotransferases
SYBR	SYBR Green I (DNA binding dye)

T

TBILI	Total bilirubin
TBX3	T-box transcription factor 3
TCA	Trichloroacetic acid
TEA	Tris-acetate-EDTA buffer
TFA	Trifluoroacetic acid
TGX	Tris-Glycine eXtended
Tm	Melting temperature
TMT	Tandem mass tag / tagging
Tris	Tris(hydroxymethyl)aminomethane
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride

U

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

Z

ZFNs	Zinc finger nucleases

Units and Miscellaneous Terms

2D	Two dimensional
3D	Three dimensional
%	Percentage
α	Alpha
β	Beta
y	Gamma
µ	Micro
n	Nano
Å	Angstrom
°C	Degrees Celsius
ΔΔCT	Delta Delta Ct
cm	Centimetres
Da	Dalton
eV	Collision energy
g	Relative centrifugal force
g	Grams
kDa	Kilo-Dalton
M	Molar (moles per litre)
mg/ml	Milligrams per millilitre
ml	Millilitre
ml/minute	Millilitres 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 ratio
µl	Microliter
µl/minute	Microliter per minute
µm	Micrometre
µmole	Micromole
µM	Micromolar
ng	Nanograms
nm	Nanometre
ppm	Parts per million
v/v	Volume per volume (millilitre per 100 millilitre)
V	Volts

Study Overview

Research and development (R&D) required to bring a single new drug to market is an intensive, laborious, multi-billion dollar undertaking for the pharmaceutical industry. Drug R&D is subdivided into pre-discovery, early and late drug discovery, preclinical evaluation, clinical trials, regulatory review, large-scale manufacturing and post-marketing. Discontinuation of new chemical entities (NCEs) or investigational new drugs (INDs) occurs within any of these developmental phases with the associated attrition rates during preclinical and clinical phases being at its highest in decades.

In an era of rapidly developing multi-drug resistance and targeted therapy, INDs need to be comprehensively tested for both efficacy as well as safety within narrow time constraints. Hepatotoxicity is a contributor to IND application failure and a major rationale for post-marketing withdrawal despite specific *in vitro* testing for this during preclinical evaluations. Early and accurate identification of toxicity is essential for success in pharmaceutical endeavours. Ethical considerations promoting the reduction, replacement and refinement (3R's principal) of *in vivo* models exacerbates the pressure faced by preclinical research. Therefore, reliable and highly predictive *in vitro* models where human derived tissue with functionality mirroring the primary tissue are required. Model-based predictions serve two important goals, to test scientific understanding and to guide decision making.

In vitro preclinical hepatotoxicity is confounded by the diverse origins of human derived tissue and the retention of hepatocellular function. Primary human hepatocytes (PHHs) are considered the "*gold standard*" for hepatotoxicity testing. However, limited availability, variation in isolation methods and inter-individual variation often preclude routine use of PHHs. Alternative tissue, especially immortalized hepatic cell lines, have become dominant in preclinical hepatotoxicity testing despite their own inherent limitations. Assessing the complete proteome of various hepatic cell phenotypes, classified according to culture technique and drug exposure should aid in determining the predictive capacity of hepatocytes in preclinical hepatotoxicity.

Chapter 1: Literature Review

1.1 Introduction

Pharmacology and toxicology are entwined disciplines, a duality described as two sides of the same coin. Paracelsus pioneered the concepts of toxicology in the 16th-century with the philosophy that all compounds possess inherent toxicity and that the dose determines safety (Niles AL, Moravec RA *et al.*, 2009). Contemporary concepts still emphasise that "*what kills can cure"* and "*what cures can kill*". Merging modern philosophies with the "Paracelsus doctrine" supports the evaluation of safety thresholds in drug R&D. Strict regulatory requirements for pharmaceuticals have resulted in the emergence of safety pharmacology. Safety pharmacology, defined by the International Conference on Harmonisation (ICH), investigates potential undesirable pharmacodynamic effects. These are secondary unwanted pharmacological or direct toxicological effects of a substance occurring in relation to exposure within or above the therapeutic range (Kinter LB and Valentin JP, 2002; Redfern WS, Wakefield ID *et al.*, 2002).

1.2 General Drug Discovery and Approval Process

Preclinical strategies, used by industry and academia alike, are designed to protect the public from harmful and inefficacious drugs. New active moieties or INDs undergo various investigational stages including pre-discovery, early and late drug discovery, preclinical evaluation, various clinical trials (Phases 0 - 4), regulatory review, large-scale manufacturing and post-marketing assessment for adverse effects (Ballet F, 1997). The viability of drug development is co-dependent on marketing rate, reasons for termination and developmental phase of IND abandonment (Kola I and Landis J, 2004).

Industry evaluation suggests that the pharmaceutical golden era is over, with very low successful drug approval rates despite increasing research niches (Cai J, DeLaForest A *et al.*, 2012; DiMasi JA, Hansen RW *et al.*, 2003). Preclinical screening from drug development to final registration can take up to 15 years. However, only one in ten molecules entering phase 1 clinical trials successfully reaches the market. It was estimated that research initiated in 2001, to undergo planned approval in 2013, would have cost on average 1.9 billion US dollars (DiMasi JA, Hansen RW *et al.*, 2003). This is a noteworthy monetary burden considering unsatisfactory approval rates. Unprecedented challenges faced by the pharmaceutical industry include high termination during clinical trials and pre-launch attrition rates, patent expiry as well as the increasing number of strict regulatory and governing policies (Cai J, DeLaForest A *et al.*, 2012; DiMasi JA, Hansen RW *et al.*, 2003).

Despite the implementation of clinical trials, patient selection and the short duration of these trials can hinder the detection of adverse drug reactions (Lasser KE, Allen PD *et al.*, 2002). Adverse drug reactions (ADRs) can be divided into: 1) dose-dependent, predictable reactions, 2) idiosyncratic responses which are not predictable or dose-related, 3) long term adaptive changes, 4) carcinogenicity and teratogenicity and 5) rebound or delayed effects (Chadeau-Hyam M, Campanella G *et al.*, 2013; Redfern WS, Wakefield ID *et al.*, 2002). This multitude of ADRs suggests the requirement for innovation by the pharmaceutical industry when assessing efficiency and effectiveness in order to match previous accomplishments in approval rates (Kola I and Landis J, 2004).

Important considerations in regulatory decisions hinge on the population risk-benefit ratio promoting positive clinical outcomes. However, decisions are based on population-level information with the knowledge that patient-level information differs in both beneficial response and susceptibility risk. Patient differences in risk-benefit ratio profiling is termed the "*efficacy-effectiveness gap*". Whereby efficacy ("*can it work?*') is the extent to which an intervention can do more good than harm under ideal clinical trial conditions and effectiveness ("*does it work?*") is the same effect within the context of the typical health care environment (Benevento M and Munoz J, 2012).

1.2.1 Drug Attrition and Drug Safety Approaches

Primary causes of drug attrition have evolved over time. In 2000, attrition was attributed equally to lack of efficacy (30%) alongside toxicology and clinical safety concerns (30%). Despite strict regulations which promote success from first-in-man trials through to drug registration (Kola I and Landis J, 2004), the pharmaceutical industry suffers high drug withdrawal rates. Highly publicised drug withdrawals span over decades from Iproniazid in 1959 to the more recent Bromfenac (1998), Troglitazone (2000) and Rofecoxib (2004)(Peters TS, 2005). Lasser *et al.* estimated that 548 new drug applications were approved between 1975 and 1999, with 10.2% (56) of these drugs acquiring black box warnings (fatal, life-threatening or disabling side effects) and 2.9% (16) being withdrawn from the market in the same period (Lasser KE, Allen PD *et al.*, 2002).

Studies assessing the physicochemical molecular properties of an IND and target-based or secondary pharmacodynamics are imperative to combat attrition rates. Identification of undesirable pharmacodynamic effects falls under the international guidelines for Safety Pharmacology Studies for Human Pharmaceuticals (ICH-S7A and ICH-S7B). Here within, the central nervous, respiratory and cardiovascular systems are specified as vital organ systems that should undergo toxicity assessments and are collectively known as the "*core battery*" of toxicity tests (Ahuja V and Sharma S, 2014; Redfern WS, Wakefield ID *et al.*, 2002).

While supplementary studies into other organ systems are conducted these are based on the nature of the candidate drug. In doing so, these guidelines less rigorously attend to organ systems such as the liver, which is a noteworthy burden for the pharmaceutical industry. Despite safety regulations being met, descriptive toxicology does not guarantee safe, efficacious medicines. Deeper exploration into the molecular and cellular underpinnings of toxicological findings is becoming increasingly essential to adequately prove biological relevance (Ahuja V and Sharma S, 2014). The interplay of predictivity, relevance and translatability (Figure 1.1) needs to be considered during the establishment of *in vitro* models used in safety pharmacology assessments (Roth A and Singer T, 2014).

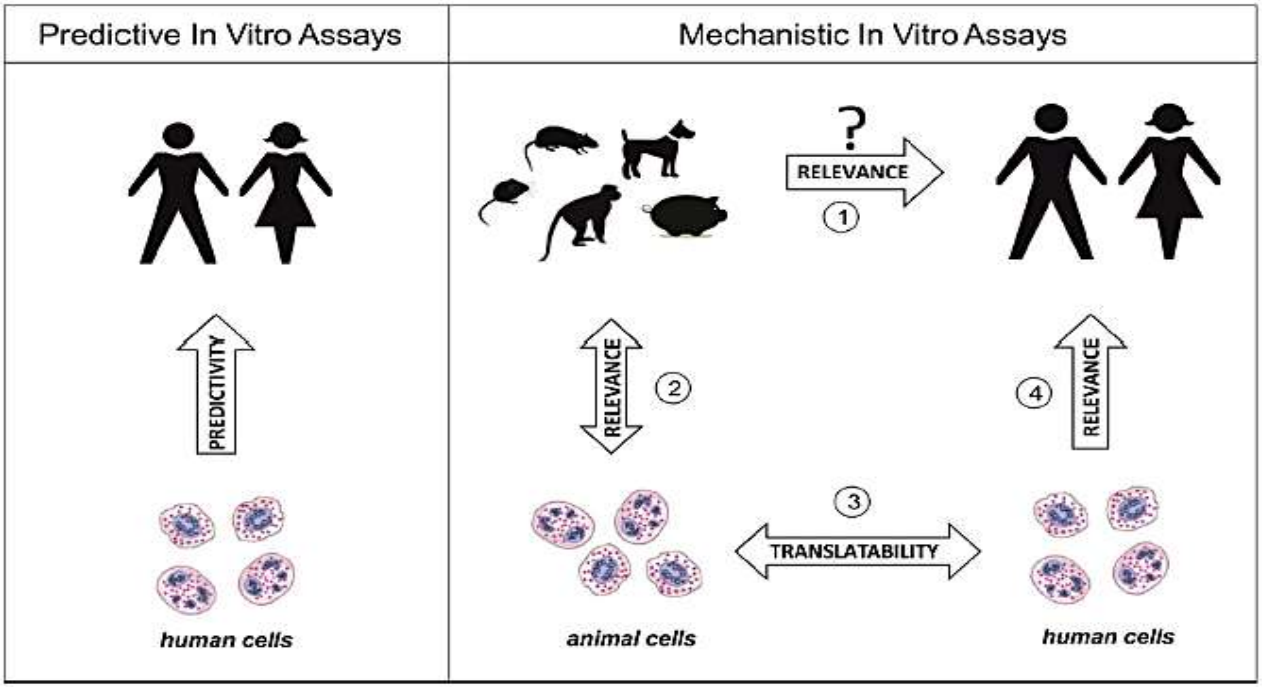

Figure 1.1: In vitro approaches for drug safety assessment: cellular assays directly predictive for the human situation (left) or a parallelogram approach (right). Potential human relevance (1) of an in vivo finding is first recapitulated in an in vitro system (2). If successful, transfer of the effect into a human cell system (3) helps in assessing human risk (4). (Reprinted with permission from Elsevier, license number 3961330921681, (Roth A and Singer T, 2014))

1.3 Safety Pharmacology

By its definition, safety pharmacology investigates undesirable pharmacodynamic effects on physiological function. Safety pharmacology is perceived as a toxicological hurdle to drug approval as opposed to enhancing safety profiles and promoting market longevity. Failing early may reduce unrecoverable cost but is not synonymous with increasing success of candidate drugs, however active pursuit of safety inevitably leads to success in the market (Feng Y, De Franceschi G *et al.*, 2014).

Streamlining safety pharmacology requires the identification of functional and molecular biomarkers of toxicity. Biomarkers are defined as a physical sign or measurement which when objectively measured can be an indicator of normal (or altered) biological processes or pharmacologic responses to a therapeutic intervention. While research to identify and validate definitive biomarkers is slow, numerous applications and insight into the mechanism of action of potential alternative drug targets are possible in this research area. Systems pharmacology, encompassing computational and mathematical modelling of complex biological systems, is defining an era of informatics-supplemented health care. Drug development models which identify clinical biomarkers and link system networks that are influenced by a particular drug, are integral to the future of safety pharmacology (Qattan AT, Mulvey C *et al.*, 2010).

1.4 The Liver

The lobed and glandular liver, the largest internal organ, is situated in the abdomen and performs many metabolic processes involving endogenous and xenobiotic compounds. The liver exhibits cellular diversity, comprising of highly structured and organised parenchymal and non-parenchymal cells (Figure 1.2). Parenchymal cells, accounting for 60% of the total liver cell population, are hepatocytes that play a central role in biotransformation, intermediary and energetic metabolism of xenobiotics. Non-parenchymal cells include liver sinusoidal epithelial cells, bile duct epithelial cells, hepatic stellate cells, Kupffer cells or liver resident macrophages and liver associated natural killer cells or pit cells (LeCluyse EL, Witek RP *et al.*, 2012).

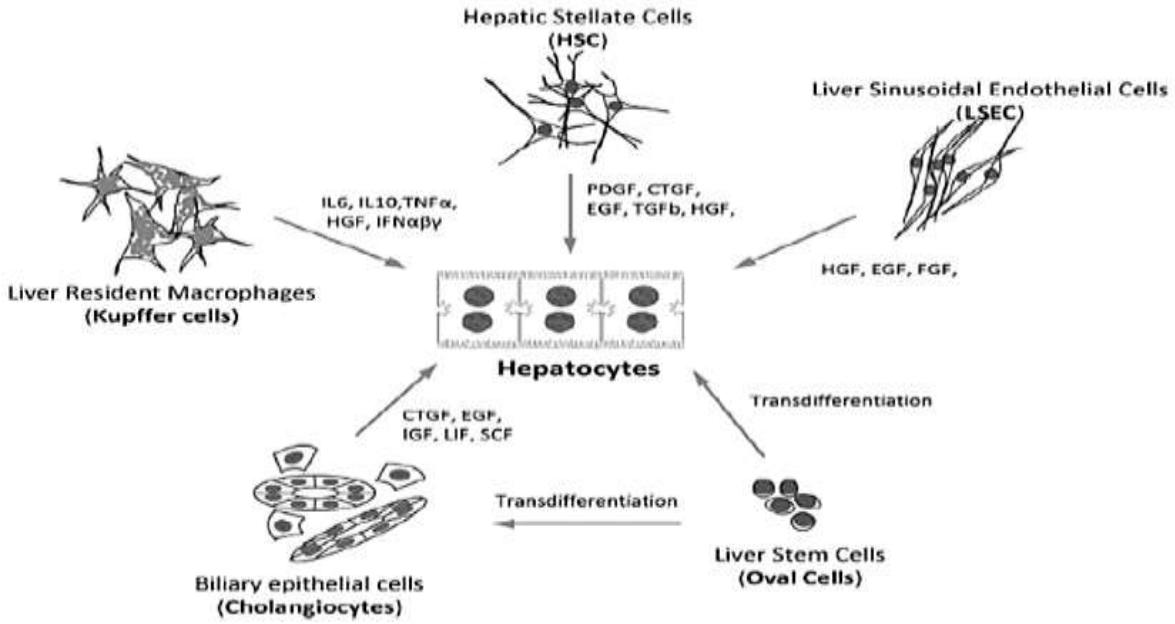

Figure 1.2: Representative cell types of the liver and their corresponding autocrine and paracrine signals that are secreted in both health and disease. (Reprinted from open access journal with permission from Taylor & Francis, (LeCluyse EL, Witek RP *et al.*, 2012))

Liver metabolism can be divided into phase I, phase II and phase III metabolism. Detoxification relies on the selectivity and abundance of specific metabolising enzymes present in liver cells. Detoxification efficiency also varies based on the substrates chemical structure, availability of co-factors, substrate selectivity, enzyme expression and localization. Phase I enzymes include cytochrome P450 enzymes (CYP450s), monoamine oxidases, flavin-containing oxygenases, amidases and esterases. These enzymes primarily catalyse biotransformation reactions resulting in the introduction of polar reactive groups (Woolf TF and Jordan RA, 1987).

The CYP450 family is a membrane-associated, hydrophobic enzyme system consisting of 57 functional genes products which carries the brunt of phase I metabolism. Subdivision into 18 families and 42 subfamilies is reflective of evolutionary relationships and is based on the percentage of amino acid sequence identity (Achour B, Russell MR *et al.*, 2014; Woolf TF and Jordan RA, 1987). Enzymes sharing more than 40% identity are assigned to a family and those sharing more than 55% identity comprise a subfamily. Gene families possess allelic variants which influence function and promote inter-individual and species pharmacogenetic heterogeneity. Major CYP450 families participating in metabolizing exogenous substances are CYP1, CYP2 and CYP3 and to a lesser degree, CYP4 (Achour B, Russell MR *et al.*, 2014; Woolf TF and Jordan RA, 1987).

Phase II enzymes include uridine 5'-diphospho-(UDP)-glucuronosyl transferase (UGTs), sulfotransferases (STs), N-acetyl transferases, methyl transferases and glutathione S-transferases (GSTs). Conjugation produces more water soluble metabolites for excretion, generally by the kidneys. Transport of phase II products across membranes, requires phase III metabolism which make use of various transporters. Transporters primarily include those facilitating uptake of organic anions and cations such as organic anion–transporting polypeptides (OATPs) and those facilitating efflux such as multidrug resistance proteins and P-glycoproteins (Woolf TF and Jordan RA, 1987). Hepatic responses to drug exposure are therefore based on the interplay of these sequentially co-ordinated metabolizing enzymes. The nature and complexity of the liver is reflected in the inability to accurately model this organ system without a high degree of extrapolation.

1.5 Hepatotoxicity

Hepatotoxicity, including drug-induced liver injury (DILI), is a leading cause of acute liver failure. DILI inflicts various patterns of injury including jaundice, hyperbilirubinemia, increases in serum levels of liver aminotransferases, cholestasis, steatohepatitis, steatosis, fatal hepatotoxicity, acute liver failure and liver necrosis. This complexity suggests that DILI will continue to present both clinical and regulatory complications for as long as NCEs and INDs enter the market (Chang C and Schiano T, 2007).

Furthermore, it is widely cited that of the approximately 10 000 documented human therapeutic drugs, more than 1 000 can be or are presently associated with DILI (Fox D, Morris L *et al.*, 2009). Chen *et al.* proposed a classification scheme whereby marketed therapeutic classes are categorized according to their DILI potential. Drug labelling classifications including "*black box warnings*", "*warnings and precaution*"' and "*adverse reactions*" are based on causality, severity and number of incidences seen in both clinical trials and post-marketing surveillance (Chen M, Vijay V *et al.*, 2011).

The number of drugs resulting in DILI remains high because there is no "*universal*" approach for early identification of hepatotoxic potential (Ballet F, 1997). Previously, the concordance between clinical pathology in humans and laboratory animals has been seen to be as low as 40%. This signifies that translation to higher integrated systems is limited and that one problem, one test paradigms are insufficient (Olson H, Betton G *et al.*, 2000).

1.5.1 Hepatotoxicity Mechanisms

Hepatotoxicity is broadly delineated in accordance with hepatotoxic potential of compounds, pattern of liver abnormalities (hepatocellular, cholestatic or mixed) or whether the mechanism of toxicity is derived from the parent or intermediate metabolite (Abboud G and Kaplowitz N, 2007; Chang C and Schiano T, 2007). Furthermore, intrinsic hepatotoxic effects are divided into direct effects at the organelle-level and indirect effects that alter metabolic pathways, cell membrane receptors and nucleic acids (Peters TS, 2005).

Mechanism-based assessment of hepatotoxicity for drugs possessing unequivocal toxicity has been outlined in the LIINTOP project [European Framework Programmes]. There is a strong reliance on multiple disruptive mechanisms in initiating hepatotoxicity. Mechanisms include, but are not limited to, bioactivation, mitochondrial impairment, oxidative stress, steatosis, phospholipidosis, cholestasis and apoptosis. Biotransformation to form excretable metabolites plays a major role in detoxification. However, biotransformed metabolites can possess intrinsic chemical reactivity promoting hepatotoxicity (Gómez-Lechón MJ, Tolosa L *et al.*, 2010). Definitive classification of the mechanism of hepatotoxicity of drugs may not provide an overview of hepatic injury, and be limited, especially due to the multifactorial nature of hepatic injury (Figure 1.3).

In addition to well described mechanisms of hepatotoxicity, idiosyncratic toxicity results from critical and complex interactions of multiple discrete genetic and non-genetic factors. These interactions are not mimicked by standard preclinical toxicity studies and occur due to a host-dependent immunologically mediated component completely independent of pharmacologic activity.

Characteristics of idiosyncratic toxicity include: 1) occurrence in less than 5% of cases, 2) inability to demonstrate a dose-response relationship, 3) injury being unrelated to pharmacological effects and 4) injury presenting with an inconsistent progressive pattern (Peters TS, 2005; Zimmerman H, 1999).

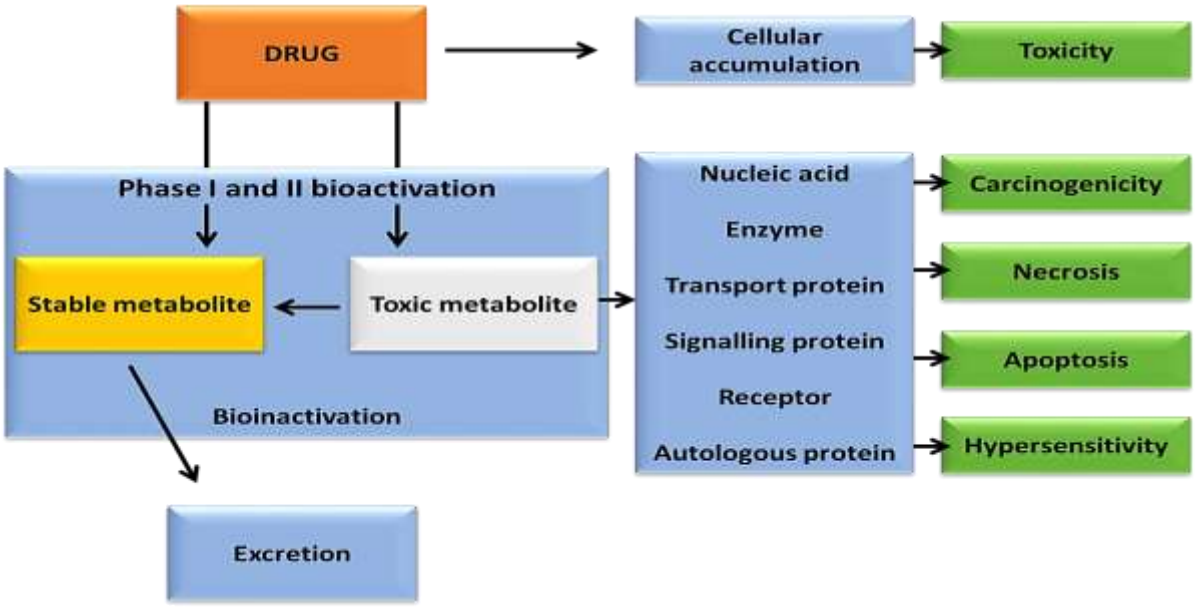

Figure 1.3: Relationship between drug metabolism and toxicity. Toxicity may accrue through accumulation of parent drug or via metabolic activation through formation of a chemically reactive metabolite. Reactive metabolites, if not detoxified, can effect covalent modification of biological macromolecules. The identity of the target macromolecule and the functional consequence of its modification will dictate the resulting toxicological response. (Reprinted/redrawn from open access journal with permission from Annual Reviews, (Park BK, Kitteringham NR *et al.*, 2005))

1.5.2 Model Systems of Hepatotoxicity

In vitro and *in vivo* studies attempt to identify cellular or molecular mechanisms that lead to toxicity as well as organ and inter-organ effects respectively. Overlapping of the results from different tests enhances the elucidation of DILI. Beyond liver cellular diversity, confounders of preclinical hepatotoxicity models include biophysical and biochemical extracellular matrix properties which influence behaviour, structure and phenotypic gene expression (LeCluyse EL, Witek RP *et al.*, 2012).

In vitro hepatotoxicity model systems (Table 1.1) include: 1) subcellular fractions including microsomes or S9 fractions, 2) primary cell cultures with a full complement of phase I and phase II metabolizing enzymes, 3) immortalized cell lines known to express reduced liver specific proteins, 4) *ex vivo* tissue slices which maintain functional heterogeneity as well as three-dimensional architecture and biochemical capacity, 5) isolated perfused organs, 6) genetically engineered cells expressing specific metabolizing enzymes and 7) non-living systems or simulations (Davila JC, Rodriguez RJ *et al.*, 1998).

Table 1.1: Advantages and limitations of *in vitro* models of hepatotoxicity

(Costa A, Sarmento B *et al.*, 2014; Davila JC, Rodriguez RJ *et al.*, 1998; Guillouzo A, Morel F *et al.*, 1997)

Model	Advantages	Limitations
Cell fractions (microsomes)	• Express phase I and phase II enzymes • Covalent binding • Inhibition or induction studies	• Loss of cytosolic enzymes • Loss of cell integrity • Short lifespan (1 - 2 hours)
Cell fractions (mitochondrial)	• Effects of drugs on respiration including ATP synthesis and β-oxidation of fatty acids	• Short term studies
Genetically engineered cells (yeasts, bacteria, non-hepatic mammalian cells)	• Express one or several hepatic functions • Covalent binding • Inhibition or induction studies • Unlimited accessibility	• Suitable for highly specific questions only (identification of cytochromes involved in the formation of toxic metabolites)
Hepatic cell lines	• Immortalized cells have limitless replication potential • Express limited liver function • Reduced cost	• Express only a subset of hepatic functions • Phenotypic drift occurs over time • Reduced expression of metabolic enzymes
Isolated hepatocytes (in suspension)	• Cells obtained from wedge biopsies • Mimic functional activity *in vivo* • Cryopreservation potential	• Limited survival (2 - 4 hours) • No bile canaliculi • Dedifferentiation and phenotypic changes
Isolated hepatocytes (in culture)	• Express functional activity for several days • Inhibition or induction studies • Interspecies studies	• Limited survival and cell growth • Dedifferentiation and phenotypic changes
Tissue slices	• Lobular structure is preserved • Human liver biopsies are suitable material • Can be applied to different species	• Preservation of functional activity is limited (less than 10 days) • Cells are not equally functional
Isolated perfused organs	• Three-dimensional structure is preserved to mimic *in vivo* environment • No interference with other structures • Functional bile canaliculi • Short term kinetic studies	• Difficult to handle • Human tissue is highly inaccessible • Short term preservation of functional activity

Ethical, economical, logistical and practical considerations dictate the choice of hepatotoxicity model. Additionally, the relationship between complexity, cost and ability to predict human toxicity have noteworthy implications. *In vitro* tools for assessing drug metabolism and safety pharmacology need to be sufficiently complex and reproducible to extrapolate to *in vivo* counterparts (Figure 1.4).

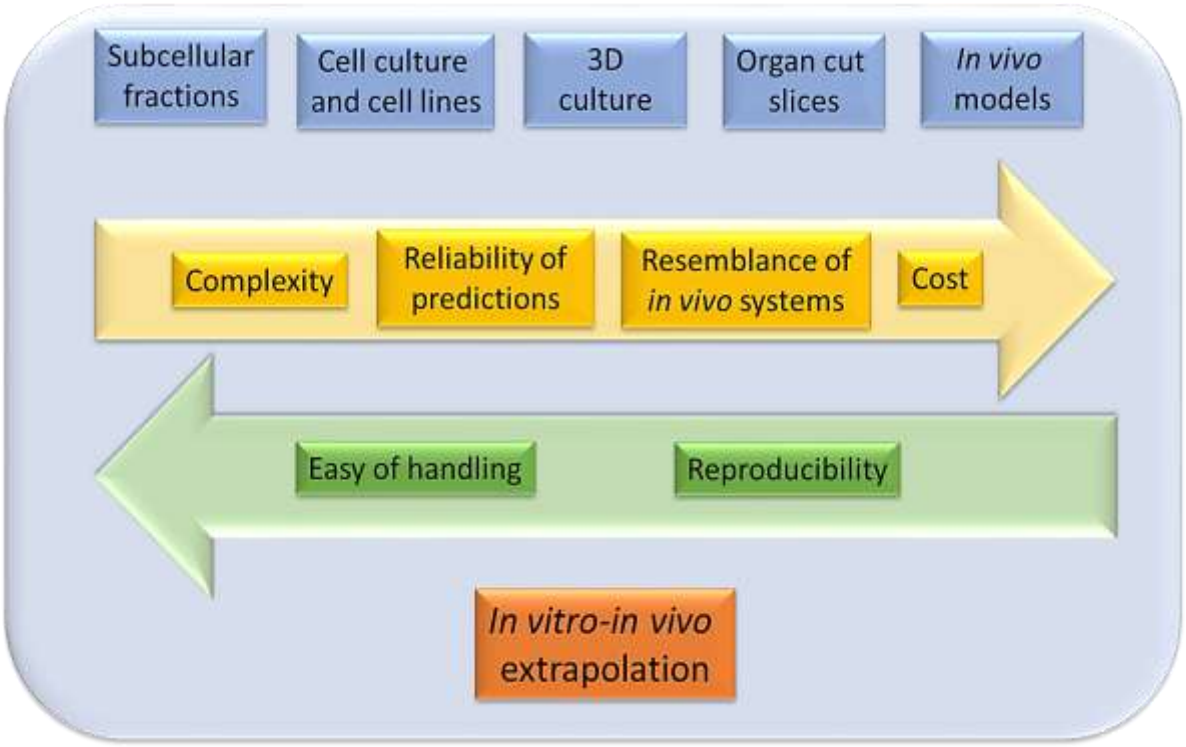

Figure 1.4: Types of model systems used preclinically to minimize in vitro-in vivo extrapolation and the features which influence them. (Reprinted/redrawn from open access journal with permission from Taylor & Francis, (Costa A, Sarmento B *et al.*, 2014))

Upon review of non-clinical and regulatory toxicity studies, the European Medicines Agency proposed an integrated risk assessment panel to tackle the multidimensional aspects of DILI. The *in vitro* hepatic liability panel includes CYP inhibition or transformed left liver lobe epithelial cell toxicity, HepG2 cytotoxicity in galactose versus glucose medium, determining inhibition of human bile salt export pumps (BSEP), steatosis and phospholipidosis assays (Ahuja V and Sharma S, 2014).

Biologically relevant *in vitro* human model systems require on-going integration with computational approaches to improve DILI prediction. The Hamner System for Health Sciences (North Carolina; USA) has developed the DILI-sum initiative, a computational model system which includes responses from multiple liver cell populations used to predict DILI. Understandably combinatorial models are not definitive enough that classifying a NCE as "*not unsafe*" is not synonymous with classifying it as "*safe*".

1.5.3 Clinical Evaluation of Hepatotoxicity

While unsatisfactory in terms of sensitivity and specificity, serum parameters are the cornerstone of clinical evaluation of hepatotoxicity. Serum markers used in the assessment of hepatocellular and hepatobiliary injury include alanine aminotransferase (ALT), aspartate aminotransferase (AST), sorbitol dehydrogenase (SDH), glutamate dehydrogenase (GDH), alkaline phosphatase (ALP), y-glutamyltransferase (GGT), 5-nucleotidase and total bilirubin. Historically, ALT served as a marker for hepatic damage due to its high abundance in the liver. However, while ALT is more liver specific than AST, it is known to increase significantly without clinical consequence (Yang X, Salminen WF *et al.*, 2012).

Currently the most successful clinical predictor of hepatotoxicity (not exclusive to DILI) is based on Hy's law. Hy's law is met when ALT or AST are elevated to equal to or greater than 3 times the upper limit of normal (ULN) and total bilirubin is elevated to 2 times the ULN (Chang C and Schiano T, 2007). Some drug classes such as anti-retroviral or anti-neoplastic drugs possess acceptable toxicity thresholds despite exceeding Hy's law criteria (Temple R, 2006). ALT and AST are inaccurate predictors because they measure the rate of cellular injury and not cellular functional capacity (Senior J, 2009) which when coupled to very high variability can results in overly cautious decision making. This necessitates new markers which predict the prognostic course, outperform serum transaminases, have strong organ specificity, correlate with histopathology and are measurable by high-throughput screening (Yang X, Salminen WF *et al.*, 2012).

1.5.4 Hepatotoxicity Biomarkers

Clinical biomarker research, which includes identification, prioritization and preliminary qualification, is painstakingly slow (Ahuja V and Sharma S, 2014). Biomarker identification in biological systems relies on integrative functions using a combination of biology, physiology, pathophysiology, ''omics" platforms and bioinformatics. Biomarkers then require qualification and validation to evidentiary and definitively be linked to specific biological processes (Fox D, Morris L *et al.*, 2009).

Predicting idiosyncratic hepatotoxicity risk has resulted in the proposal of novel mechanistic biomarkers. A "*bedside to bench"* approach is encouraged to improve risk evaluation, whereby clinical findings are translated into new preclinical models. Novel biomarkers, in a panel based approach, have provided information about the multicellular and multi-mechanistic basis of acetaminophen induced hepatotoxicity in clinical patient stratification. As a descriptive tool, use of biomarker panels is a potential approach for preclinical research whereby multiple biomarkers can stratify "*go/no-go*" decisions for specific drug classes (Fox D, Morris L *et al.*, 2009).

1.6 Sources for *In Vitro* Hepatotoxicity Models

Cell-based hepatotoxicity screening is essential prior to further R&D commitment. Predictability varies based on cellular functionality, state of hepatocyte differentiation, duration of exposure and the investigational drug (Knasmüller S, Mersch-Sundermann V *et al.*, 2004; Xu JJ, Diaz D *et al.*, 2004). Elucidating the variance in hepatotoxicity models is fundamental to understanding the biological implications.

1.6.1 Primary Cells

Freshly isolated primary human hepatocytes (PHHs) are the *"gold standard"* or "*historical standard*" for functional hepatocytes (Mills JB, Rose KA *et al.*, 2004). While restricted in their lifespan and unable to proliferate, under the appropriate conditions PHHs express all major metabolising enzymes and transporter proteins. However, rapidly decreasing metabolic function and liver-specific gene expression occur in culture (LeCluyse EL, Witek RP *et al.*, 2012). Commercially available PHHs include plateable or suspension formats, single or pooled donors, fresh or cryopreserved cells which can be induction, transporter or metabolism qualified. However, donor-dependent variation still limits population-level extrapolation (Mills JB, Rose KA *et al.*, 2004). Additionally, formats mimicking the total liver cell population with stellate, progenitor and intra-hepatic biliary epithelial primary cells are available. Newly available variants of PHHs (upcyte® technology by Medicyte, Heidelberg; Germany) have been designed to overcome the inability of these cells to proliferate using a viral gene transfer system. Variations in PHHs alone, further exemplifies the concerns of modelling *in vitro* hepatotoxicity.

1.6.2 Immortalized Cell Lines

Immortalized cells proliferate indefinitely and are produced by the cellular nature of oncogenic transformation or transfection. These are generally considered acceptable for hepatotoxicity testing provided that CYP1A2 and CYP3A4 induction can be demonstrated. One limitation with immortalized cell lines is the tendency to change phenotypically with each passage number due to culturing in relatively well defined media. Immortalized cell lines include human hepatocellular carcinoma HepG2 cells, Fa2N4 cells, bipotent hepatic progenitor HepaRG cells and HepG2 clonal variants which have increased albumin and alpha-fetoprotein expression such as HepG2/C3A cells. Due to ease of use and replication potential, HepG2 cells are the most commonly used and characterized for *in vitro* hepatotoxicity. However, HepG2 cells have been associated with reduced basal gene expression of phase I and phase II biotransformation enzymes, discontinuous phenotypic and functional properties as well as genetic instability (Duret C, Gerbal-Chaloin S *et al.*, 2007; LeCluyse EL, Witek RP *et al.*, 2012).

1.6.3 Stem Cells

Stem cells are undifferentiated cells which occupy a unique position in cellular hierarchy, with the ability to self-renew and differentiate. Lineage differentiation is dependent on the stem cells potential which exists on a continuum of decreasing potencies from totipotent to unipotent. Two major classes of stem cells are embryonic stem cells (ESCs) and adult stem cells (ASCs). ASC are lineage restricted, residing in their respective tissues (mesenchymal, neuronal, umbilical, cardiac, retinal and limbal) and remain quiescent until activation via injury or disease.

ESCs form embryoid bodies which can either be directly differentiated into three germ layers (endoderm, mesoderm and ectoderm) via directed differentiation, or may give rise to teratomas (germ cell tumour containing several tissue types) through uncontrolled differentiation *in vivo*. Advances in cellular reprogramming strategies have revealed a third stem cell type, referred to as induced pluripotent stem cells (iPSCs). iPSCs are somatic cell which have been dedifferentiate to pluripotency and can be differentiated into germ layer specific cellular lineages (Singh RK, Gaikwad SM *et al.*, 2014).

1.6.3.1 Induced Pluripotent Stem Cells

The concept of altering one cell type into another originated in the work of developmental biologist Sir John Gurdon, who in 1958 pioneered nuclear transplantation. Developmental biology has undergone revolutionary advances with reprogramming of somatic cells to pluripotency. Reprogramming was achieved using transcription factors involved in maintenance of pluripotency and stem cell phenotypes: OCT-3/4, SOX-2, c-MYC and KLF-4. Direct reprogramming made use of stochastic overexpression of transcription factors in a process of transgenesis whereby exogenous genes promote transmissibility of new properties to the cells progeny (Takahashi K and Yamanaka S, 2006).

Successfully deriving iPSCs is governed by transcription factors, method of factor delivery (genome integrating or genome non-integrating methods), cell type, degree of reprogrammability, time required to becoming independent of exogenous factors, derivation conditions, correct identification of iPSC colonies, expansion and characterisation of morphology, molecular features and functionality (Maherali N and Hochedlinger K, 2008; Rashid ST and Vallier L, 2010). Due to the multiplicity of factors driving reprogramming, a minimal set of criteria for assessment of reprogramming has been suggested. These include morphological attributes, unlimited self-renewal, expression of pluripotency genes (key pluripotency factors: OCT-4, NANOG and ES-specific antigens: SSEA-3/-4, TRA-1-60/-81), down regulation of genes specific to the original cell lineage, transgene independence and functional differentiation (Maherali N and Hochedlinger K, 2008).

1.6.3.1.1 Opportunities for Use of iPSC Technology

Predictive toxicology is a limitation within the pharmaceutical industry. Readily accessible tissues such as skin, peripheral blood and other somatic cells enable generation of iPSC libraries using genetically diverse, population specific cells. Furthermore, patient or disease-specific iPSCs can be generated for use as *in vitro* disease models, shaping new horizons for personalized medicine (Figure 1.5). Aspects of reprogrammability are continually undergoing refinement and if standardised, will revolutionize the generation of iPSC libraries. Human iPSCs differentiated into hepatic, neuronal or cardiac lineages have applications in regenerative medicine and pharmaceutical science (Anson BD, Kolaja K *et al.*, 2011).

1.6.3.1.2 iPSC-derived Hepatocyte-like Cells

Advantages of iPSCs are that sourcing and self-renewal remain inherent while possessing the potential for differentiation into biliary and hepatic lineages. Defined soluble factors promote iPSC differentiation to definitive endoderm through hepatic specification and into hepatocyte-like cells (HLCs). However, differentiation protocols require robust characterisation. Defined mRNAs clusters have been suggested to track phenotypic changes during differentiation. Dynamically altered genes identified in the definitive establishment of hepatocytes included FOXA2, TBX3, HHEX, HNF4α, GATA4 and GATA6. However, HLCs also express mRNAs not usually associated with either foetal or adult hepatic tissues. Despite HLCs not accurately mimicking hepatic gene expression, identification of authentic regulators of differentiation aid in tailoring protocols (DeLaForest A, Nagaoka M *et al.*, 2011).

An advantage of HLCs is the ability to generate isogenic cells which are precisely engineered to reflect a specific patient population or individual. Genetic variants from cohorts of healthy, diseased and DILI patients can then be compared. However, HLCs obtained from most protocols fail to express drug metabolising enzymes at levels comparable to PHHs (Sampaziotis F, Segeritz CP *et al.*, 2014).

Despite success using numerous protocols, transcripts, protein expression and functional activity reflect foetal more than a mature adult hepatic phenotype. This brings into questions the lineage specificity of differentiation or suggests that HLCs are stuck in a foetal-like state. Factors required to transition to mature a HLC phenotype are still elusive (Baxter M, Withey S *et al.*, 2015). However, innovation in systems biology makes HLCs a useful platform for further study. The demand for reliable hepatic tissue for drug screening and other *in vitro* methodologies has resulted in various commercially available HLCs which include iCell® Hepatocytes derived from iPSCs (Cellular Dynamics International, Wisconsin; USA).

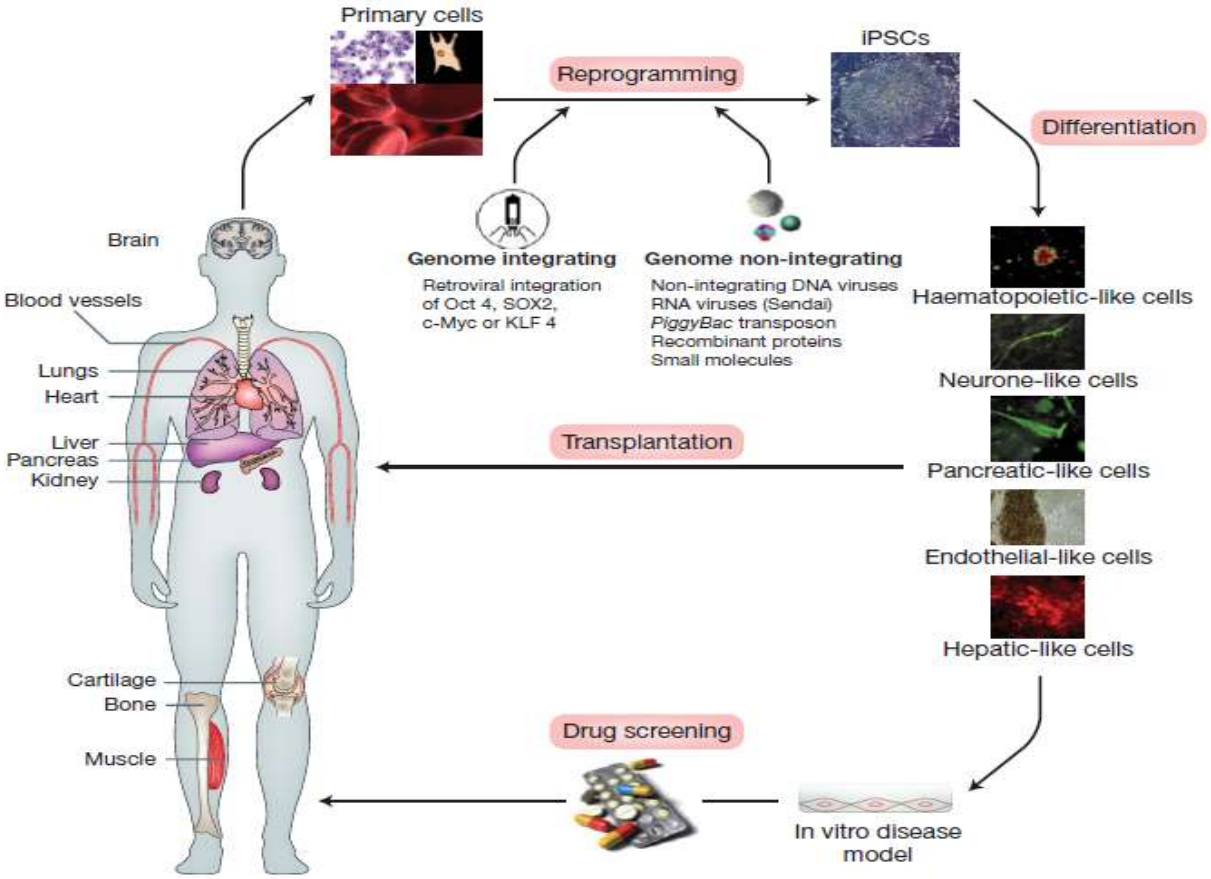

Figure 1.5: Reprogramming can be performed starting with a variety of donor somatic cells (depicted here are fibroblasts and blood progenitors). The reprogramming factors can be delivered via either genome integrating (classical approach as first used by Yamanaka which involved the integration of the reprogramming factors into the somatic cell genome) or genome non-integrating methods. Reprogramming the donor cell back to an induced pluripotent state is followed by differentiation to a target cell(s) of interest. Patient-specific cells should carry the unique genetic signature of the donor and could therefore be used for different objectives including in vitro disease modelling or drug screening. It is hoped that in the near future such reprogrammed cells may also form part of a cell-based autologous transplantation strategy for treating diseases. Examples of cells ideally suited to such a strategy include hepatic, neuronal and blood cells. (Reprinted with permission from Cambridge University Press, license number 3961350415127, (Rashid ST and Vallier L, 2010))

1.6.4 Comparison of *In Vitro* Hepatotoxicity Models

Whole genome assessment of differentiated and undifferentiated HepaRG cells, HepG2 cells, PHH and human liver tissue have been used to generate a similarity matrix of gene expression. This has shown high transcriptomic correlation between differentiated and undifferentiated HepaRG cells (0.966) as well as between liver tissue and PHH (0.920). Strong transcriptomic correlations also existed between differentiated HepaRG's and PHH (0.891) as well as differentiated HepaRG's and liver tissue (0.881). While weak correlation (0.791) exists between liver tissue and HepG2 cells (Hart SN, Li Y *et al.*, 2010).

While gene transcription levels of inducible, metabolically competent HepaRG cells express a high degree of similarity to PHHs and liver tissue at mRNA levels, this does not necessitate similarity at the proteome or enable translation into detection of hepatotoxicity (Kampf C, Mardinoglu A *et al.*, 2014; Kim M-S, Pinto SM *et al.*, 2014). Available data supports the notion that there is, as of yet, no single hepatotoxicity model that accurately reflects the diverse array of possible outcomes of different types of cellular toxicity.

Table 1.2: Comparison of cells used for *in vitro* toxicity testing
(Anson BD, Kolaja K *et al.*, 2011)

Cell source	Origin	Quantity	Genetic stability	Genetic diversity	Disease model	Gene targeting
Transformed cells	Immortalized or tumour cells	Unlimited	Aneuploid with drift	Minimal	Possible, sometimes engineered	Yes, but limited
Primary cells	Primary tissue isolation	Limited	Diploid stable	Moderate	Intrinsic mutations possible	No
ES cells	Embryo	Unlimited	Diploid, relatively stable	Moderate	For abnormal pre-implantation testing or engineered	Yes
iPS cells	Somatic cells	Unlimited	Diploid, unknown	Great	Greatest with direct sampling of diverse populations	Yes

1.7 Cell Culturing Techniques

Cellular growth conditions or the "*bio-mimetic microenvironment*" can influence adhesive cues, the extracellular matrix (ECM), growth factors, intercellular contact, mechanical forces, geometry, cell shape and spatial organisation which in turn collectively dictate cellular functionality and responses (Bhadriraju K and Chen CS, 2002). Strong reliance on the microenvironment complicates which is greater: the dynamic range of gene expression between cell types or environmentally-induced phenotypic states of a single cell type (Shin JW, Park SH *et al.*, 2012).

Use of static, two-dimensional (2D) monolayer culture techniques oversimplifies the normally more complex cellular microenvironment, diminishing gene expression and inducing phenotypic modulation. This then inappropriately represents the *in vivo* counterparts. However, *in vitro* assessments are unavoidable and manipulation of the microenvironment may enable retention of critical functionality (Bhadriraju K and Chen CS, 2002; Peters TS, 2005). Cellular dependence on "*community behaviour*" has resulted in use of three-dimensional (3D) cultures capable of spatial organisation.

Three-dimensional cultures forming spheroidal structures enable cell-cell interactions in all dimensions promoting physiological boundaries which more closely approximate *in vivo* arrangements (Fey SJ and Wrzesinski K, 2012). These cellular interactions and communication in co-existing matrices are essential in determining biological responses to drugs.

Table 1.3: Advantages and limitations of 2D monolayer and 3D spheroid cultures
(Bhadriraju K and Chen CS, 2002; Fey SJ and Wrzesinski K, 2012; Kim JB, 2005; Shin JW, Park SH *et al.*, 2012)

Culture condition	Advantages	Limitations
Monolayer cultures	• Standardised • Easy to use	• Non-natural morphology • Rapidly diminishing differentiation functions • Non-physiological effects • Lack structural architecture and stroma • Lack cell-cell and cell-matrix interactions reducing the ability to extrapolate to *in vivo* • High surface area exposure to test compounds
Spheroid cultures	• Spatially organised • Co-existing matrices approximate *in vivo* environment • Mimics physiological barriers • Co-cultures for intercellular signalling	• Inability to count cells in a spheroid • Grow slower than 2D counterparts • High cell number per volume of medium • Suboptimal nutrient-waste exchange conditions • May develop central necrosis or regions of hypoxia (mimicking cancer physiology)

Limitations of spheroid cultures are as numerous as those for 2D monolayers. However, understanding the development, growth and architecture of spheroids is less limiting than navigating diminished functional responses associated with monolayers. Platforms for development of spheroid cultures include dynamic cultures which aid in the flow of oxygen and nutrients (spinner flasks, rotating wall vessel bioreactors) and static cultures which have non-adhesive surfaces encouraging cell suspension (scaffold-supported and scaffold-free platforms, liquid overlays, magnetic levitation, hanging drops).

Three-dimensional culture formats such as soft agarose gels or hanging drops are easy to adopt and can be used to produce spheroids which aggregate over time and can be cultured extensively. Growth tracking of spheroids (Figure 1.6) illustrates cellular organisation, compaction and growth. Co-cultures of multiple cell types, in varying ratios, can be produced to mimic intercellular signalling. Co-culture formats include: 1) seeding multiple cell types together allowing spontaneous self-aggregation, 2) sequential seeding where one spheroid is formed and then encapsulated by a second cell type and 3) merging of two simultaneously formed spheroids (Mehta G, Hsiao AY *et al.*, 2012).

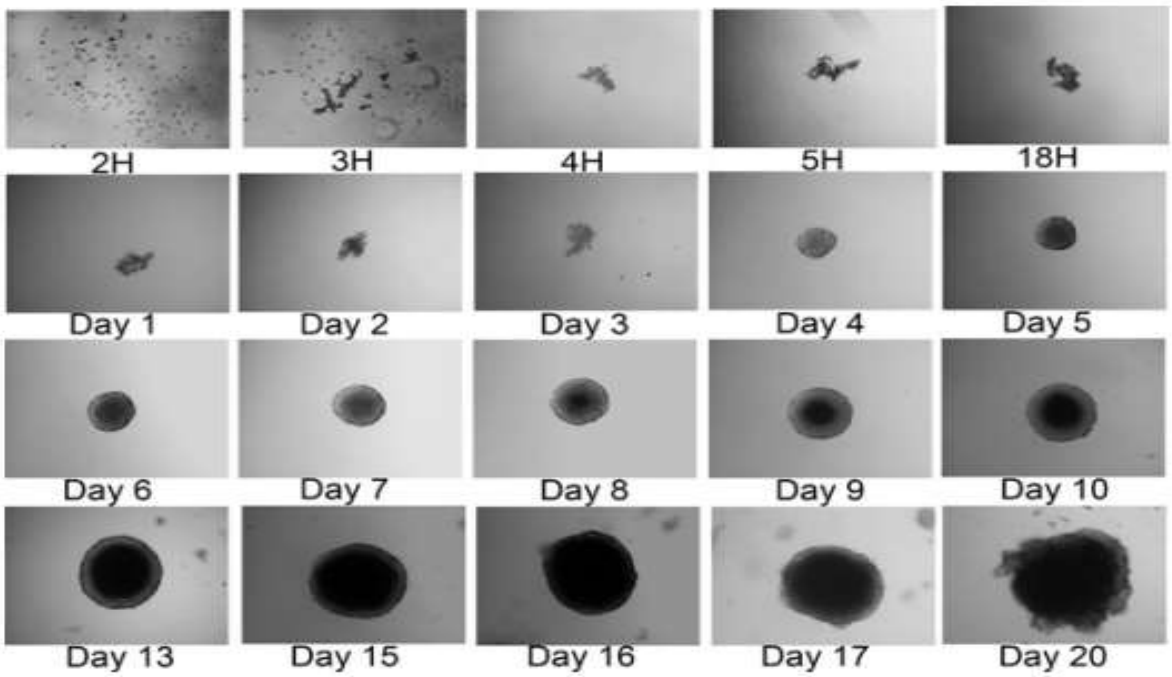

Figure 1.6: Growth tracking of human neuroendocrine tumour cell (BON-1 3D) spheroids on agarose-coated plates. Cells were imaged over 20 days with a Zeiss Axiovert 200/M-based phase-contrast microscope. (Reprinted from open access journal with permission, (Wong C, Vosburgh E *et al.*, 2012))

Complex rotary cell culture systems (RCCS) enable growth in a simulated microgravity environment generating spheroids of 1 - 3 mm within 48 hours. Central necrosis is a limitation of larger spheroids. However, in some laboratories no apoptosis has been observed in spheroids of less than 1 mm in diameter. For most cell types, growth characterization shows proliferation at the spheroid periphery, while HepG2 cells in RCCS reveal a more disorganized structure with chaotic cell death throughout (Baxter M, Withey S *et al.*, 2015; Li VC and Kirschner MW, 2014). Cell-printed constructs for tissue engineering, *in vitro* drug screening and tissue model applications may promote long-term cell survival in 3D culture. Bioprinting in a decellularized extracellular matrix bioink is used to grow 3D structured tissues with intrinsic cellular morphologies (Fox D, Morris L *et al.*, 2009).

Research application, cost, infrastructure, cell-type and origin, incorporation of physical signals, ECM and effective mimicry of native tissues (viscoelastic properties, interstitial flow and transport characteristics) govern the appropriateness of 3D techniques. Technical aspects of modelling influence cellular physiology and function and need to be considered in liver modelling. Three-dimensional models for predictive assessments of drug safety need to have biological relevance, be medium to high throughput with high standardization and provide a quantitative "*go/no-go*" decision. The many factors required to imitate organ systems make it difficult to establish which models to refine (Roth A and Singer T, 2014).

Table 1.4: Advantages and limitations of different types of 3D cultures
(3D-Biomatrix, 2013; Dhaliwal A, 2012; Kim JB, 2005)

Culture method	Advantages	Limitations
Bioreactors (spinner flask, perfusion bioreactors)	• Scaffolds: high volume cell production • Dynamic perfusion bioreactors: improves nutrient-waste exchange	• Heterogeneity in spheroid size • Scaffolds may be incompatible with downstream applications
Gels (liquid overlay cultures)	• Multiple ECM mixtures available (natural, synthetic, animal-based) • Cells do not adhere to substratum but migrate and aggregate • Hydrogels (hydrophilic polymer chain network) with tissue-like flexibility and viscoelasticity	• Animal derived ECM extracts may contain undefined constituents promoting batch-to-batch variations • Inaccurate reproducibility • Organization of cells can impart changes in gel structure
Scaffold-free platforms (hanging-drops, magnetic levitation)	• Spontaneous cell aggregation • Self-encapsulate in ECM • No support or artificial substrate • Self-sorting histotypic structure • Homotypic or heterotypic aggregation • Vary size by adjusting seeding density	• Oxygen perfusion-diffusion limits • Limited to critical diameter (500 - 600 µm) to circumvent central necrosis
Scaffolds (include microspheres)	• Natural (silk, collagen, gelatin, fibrinogen, hyaluronic acid, alginate), synthetic (PEG), mixed polymers, metals and ceramics • Customizable to different porosity, interconnectivity, permeability and mechanical (tensile strength, elastic modulus) characteristics	• Biocompatibility, wettability, mechanical properties, degradability, surface chemistry and scaffold-to-scaffold variability • Use of cyto-incompatible solvents • Membrane integrity compromised in fibrous and highly porous scaffolds • Often incompatible with downstream applications
Microchips	• Integrate microfluidics technologies • Organ specific systems • High gas permeability	• Limited commercial availability • Limited HTS • Requires expertise

1.8 "OMICS" Technologies

The *"multiple determinant hypothesis"* suggests that hepatic injury is related to a sum of factors. Systems biology aims to reconcile shortcomings of conventional highly targeted studies by employing multiple platforms which consider various collective constituents to identify relevant biologically mediated signatures of toxic exposure. This is achieved using complementary data from the unchanging genome to the dynamic transcriptome, proteome and metabolome (Au J, Navarro V *et al.*, 2011; Yang X, Salminen WF *et al.*, 2012).

1.8.1 Proteomics

The proteome is a dynamic entity, defined as the entire complement of proteins including isoforms and modifications, protein interactions, structures and complexes in both spatial and temporal forms. Proteomic analysis is hindered by the range of protein size and abundance, sample variability and degradation, developmental specificity and disease or experimental perturbations (Tyers M and Mann M, 2003). Experimentally, the core proteome describes abundant and easily detected proteins commonly seen in every replicate. Proteomics can be used to monitor all proteins present in a sample or selected proteins when the complexity and dynamic range reduce the probability of identification. Regardless of reduction in sample complexity, heterogeneity of individual proteins with various isoforms and post-translational modifications (PTMs) still complicates identification and quantitation (Piersma SR, Warmoes MO *et al.*, 2013).

Assessing biologically relevant changes requires measuring the altered proteome in relation to the original state of the system. Qualitative and quantitative proteomic approaches (Figure 1.7) include bottom-up and middle-down proteomics of enzymatically or chemically cleaved proteins (15 amino acid peptides of roughly 1.6 kDa to 20 kDa fragments) or top-down proteomics of in-tact proteins (Meyer B, Papasotiriou DG *et al.*, 2011).

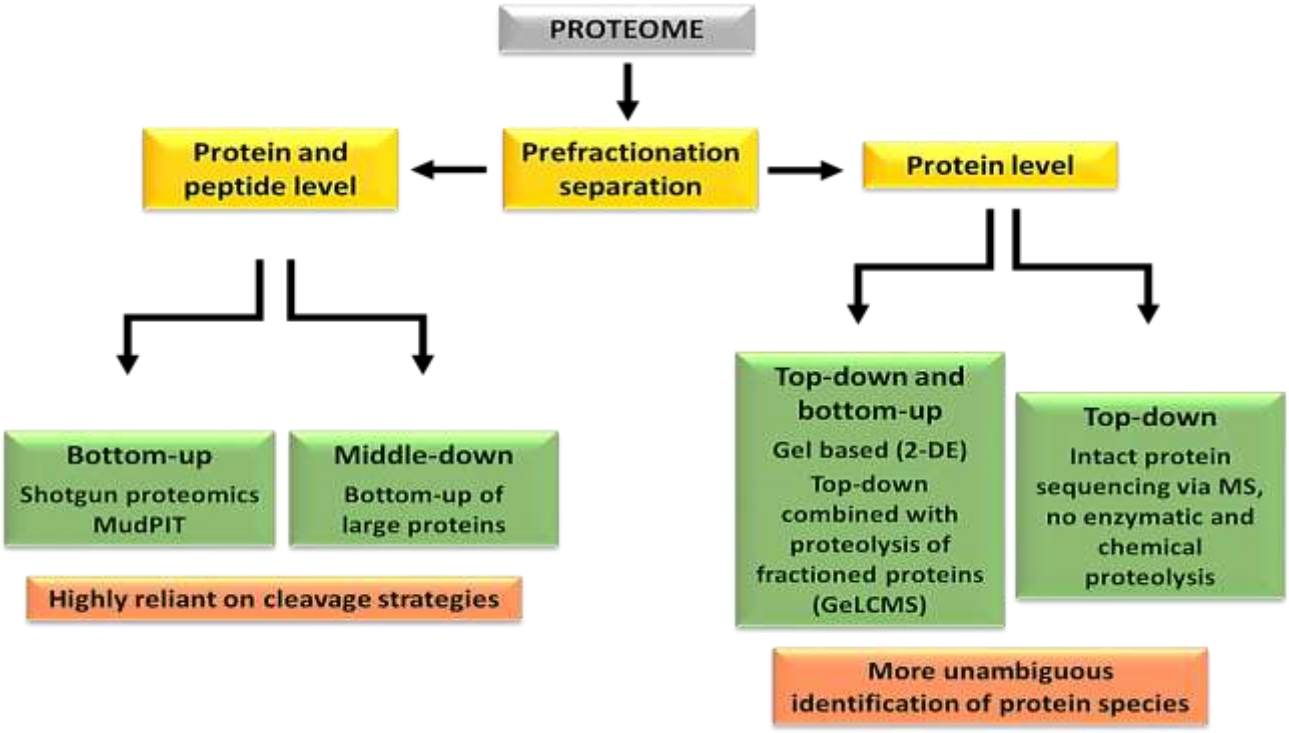

Figure 1.7: Summary of strategies used to investigate the proteome. (Reprinted/redrawn with permission from Springer, license number 3961351270809, (Meyer B, Papasotiriou DG et al., 2011))

Accurate multi-dimensional proteomics is governed by chemical, physical and biological relevance (Figure 1.8). Assessing these properties and their implication in cell biology is dictated by factors influencing proteome coverage. Prominent influencers include detergents and denaturants used for solubilisation, enzymatic reagents for bottom-up proteomics which have variable specificity and generate proteotypic signatures, instrument sensitivity, labelled versus label-free approaches, the ability to maintain PTMs and the formation of protein complexes (Larance M and Lamond AI, 2015).

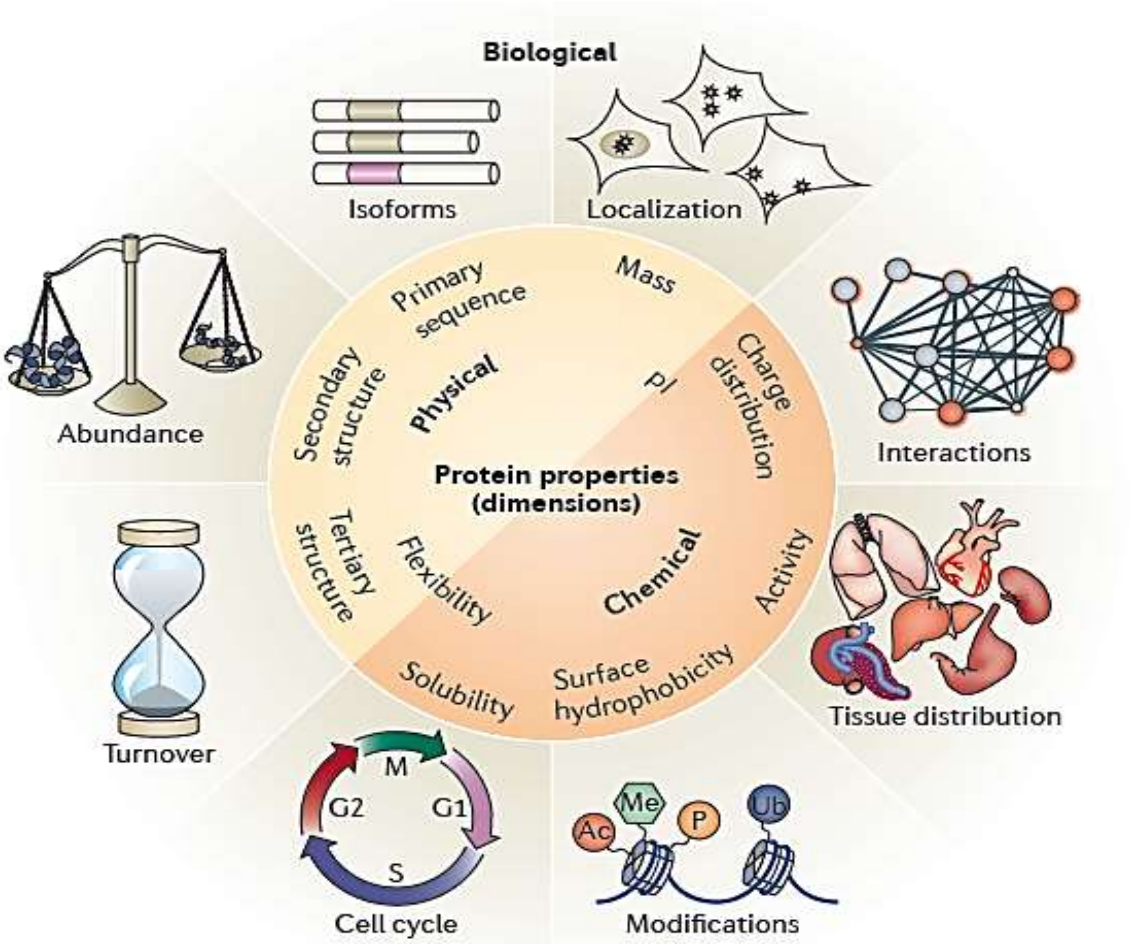

Figure 1.8: Multidimensional proteome analysis of cells and tissues. Proteins can have many different properties (dimensions) that are either largely physically (light orange shaded area), chemically (orange shaded area) or biologically (beige shaded area) relevant. This figure illustrates some properties important for cell biology research and those that need to be taken into consideration when developing new separation methods for multidimensional analysis. (Reprinted with permission from Nature Publishing Group, license number 3961360127642, (Larance M and Lamond AI, 2015))

1.8.1.1 Gel-based Proteomics

Despite exponential advances in proteomic strategies, one and two dimensional-polyacrylamide gel electrophoresis (SDS-PAGE and 2D-PAGE) remain versatile platforms for global protein separation. Gel-based proteomics can be laborious and limited by depth of coverage and dynamic range, poor detection of low abundance proteins or those with extreme characteristics (large, acidic, basic, hydrophobic), sensitivity of staining techniques, incompatibility of staining with downstream applications, gel-to-gel-variation and inadequate normalization (MacCoss MJ and Yates JR, 2001; Tyers M and Mann M, 2003).

Protein bands or spots can be excised, enzymatically digested (in-gel digestion: IGD) and the peptides identified using mass spectrometry (MS), often termed GeLCMS, which yields protein identification and quantification (MacCoss MJ and Yates JR, 2001). Hydrophobic or large peptides have limited extraction after IGD, resulting in approximately 15 - 50% loss of separated proteins. Furthermore, large proteases (25 kDa) including trypsin, pepsin and Glu-C, do not penetrate the gel matrix sufficiently and reduce cleavage efficiency (Meyer B, Papasotiriou DG *et al.*, 2011).

1.8.1.2 Mass Spectrometry-based Proteomics

Proteomics samples are usually exceedingly complex with high abundance proteins dominating detection. MS-based techniques, making use of mass-charge-ratios (m/z), has revolutionized proteome exploration. The origins of MS-based proteomics include the use of automated 2D proteome-wide technology known as multidimensional protein identification technology (MuDPIT) by John Yates and Ruedi Aebersold's quantifying complex mixtures using isotope-coded affinity tags (Gygi SP, Rist B *et al.*, 1999; Washburn MP, Wolters D *et al.*, 2001) .

Mass spectrometry provides the ability to elucidate many PTMs and other functional interactions. This is associated with improved dynamic range (approximately 10^4) and the ability to be selective specific peptides during data acquisition. For instance, the exclusion of highly abundant peptides can be done in order enhance the identification of low abundant peptides. Analysis of proteins requires peptides or proteins to be in the gas phase for fragmentation and detection. Common soft ionization methods used for proteomics are electrospray ionization (ESI) for peptide sequencing or matrix-assisted laser desorption ionization (MALDI) for whole protein and peptide mass fingerprinting. Despite success with MS methodologies most MS detection is saturated at 10 000 - 11 000 proteins, which is a limitation when dealing with protein interactions and variable protein abundance within complex samples (Altelaar A and Heck AJ, 2012; Demartini DR, 2013).

Successful MS-based proteomics is reliant on the application of appropriate techniques and sensitive instrumentation. Untargeted protein identification with relative quantitation includes label-free approaches, isobaric mass tags, stable isotope labelling by amino acids in cell culture (SILAC), dimethyl labelling and isotope-coded affinity tags (ICAT). Targeted analysis of pre-determined peptides can be quantified, relatively or absolutely, using intelligent selected-reaction monitoring (iSRM) and high-resolution, accurate mass (HR/AM) selected ion monitoring (SIM) scans (HR/AM-SIM).

Isobaric tags for relative and absolute quantitation (iTRAQ, AB Sciex) or tandem-mass tags (TMT, Thermo Scientific) were developed for bottom-up quantitative proteomics and are ideal for untargeted, simultaneous, proteome wide comparisons. TMT employs N-hydroxy-succinimide (NHS) chemistry to tag peptide N-terminal amine groups and ε-amine groups of lysine residues (Ross PL, Huang YN *et al.*, 2004; Thompson A, Schäfer J *et al.*, 2003). Isobaric tags are classified by the mass/charge (m/z) shift of the specific tag used for the chemical modification and the number of samples that can be multiplexed (Pichler P, Köcher T *et al.*, 2010; Thompson A, Schäfer J *et al.*, 2003).

Isobaric tags (Figure 1.9) have the same nominal precursor mass, and are structurally composed of an amine-reactive NHS-ester group, a spacer arm or balance and a distinguishable MS/MS reporter group for quantification. Samples are individually denatured, alkylated, trypsin digested, specifically labelled with one tag and combined at the same protein concentration for simultaneous chromatographic analysis.

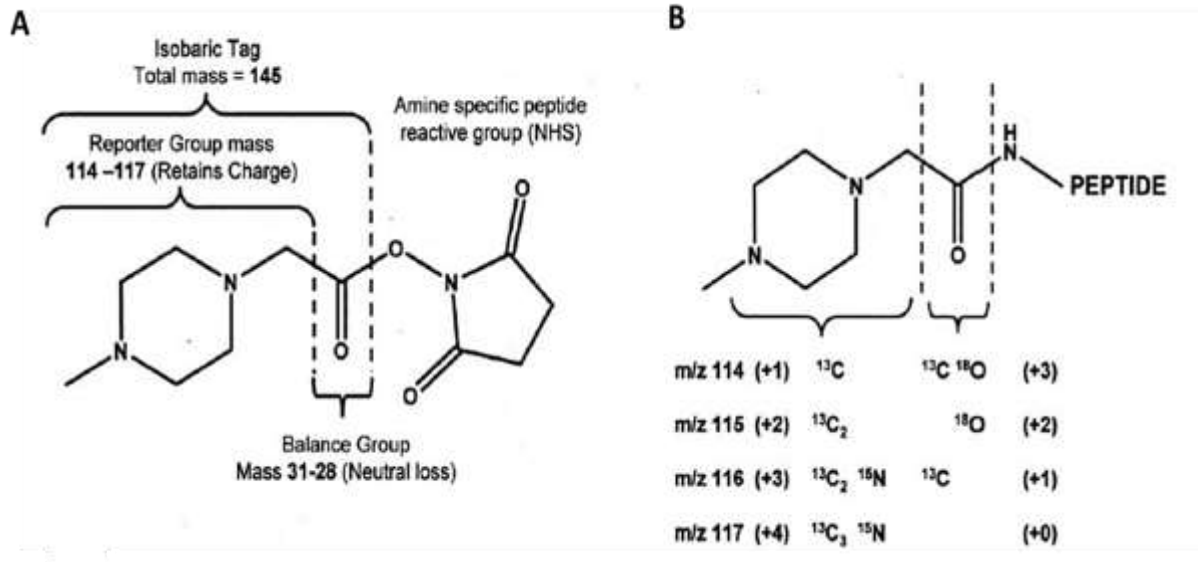

Figure 1.9: A) Diagram showing the components of the multiplexed isobaric tagging chemistry. The complete molecule consists of a reporter group (based on N-methylpiperazine), a mass balance group (carbonyl), and a peptide-reactive group (NHS ester) and B) The overall mass of reporter and balance components of the molecule are kept constant using differential isotopic enrichment with 13C, 15N, and 18O atoms. (Reprinted with permission, (Ross PL, Huang YN *et al.*, 2004))

During MS/MS analysis MS1 scans (precursor ion scans) show a single unresolved tagged precursor ion peak from each sample present in the multiplexed combination. Fragmentation releases the mass-balancing carbonyl moiety (neutral loss), the singly charged reporter ion (simultaneous analysis of up to ten multiplexed reporter ions) is used for quantification and the peptide fragment peaks of b- and y-ions are used for peptide sequence identification (Pichler P, Köcher T *et al.*, 2010).

Variability in identification and quantification can be introduced during sample preparation, by the type and quality of instrumentation, tag-peptide ratio effects from labelling variable amounts of protein, acquisition efficiency in liberating reporters and data analysis. TMT and iTRAQ can suffer distorted protein identification or quantification as a result of ratio compression caused by co-fragmentation of co-eluting peptides during acquisition. Furthermore, mechanisms regulating protein function which can be inducible, sequential or antagonistic in nature can result in reduced PTM identification when using isobaric tagging (Pottiez G, Wiederin J *et al.*, 2012).

Proteomic data, and the biological conclusions inferred, can be distorted by population averages of the cell types making up a tissue. Single cell proteomics, whereby simultaneous introduction of a digesting enzyme and sample can occur using an online system may reduce this population averaged distortion. Advances in analytical techniques have been developed for single cell proteomics to reduce issues with sample averaging (Sawyer N, Worrall L *et al.*, 2008; Yan W, Hwang D *et al.*, 2008).

1.8.1.3 Bioinformatics in Proteomics

Peptide identification requires database search strategies which match MS2 (product ion scans) spectra with theoretical and experimentally digested peptide spectra. Matched candidate peptides are scored based on mathematical and algorithmic strategies. Database searching and consolidation from various data sets is influenced by many key factors, including sample quality and complexity, instrument type, number of fractions or replicates, digestion enzymes, missed cleavages, fixed and variable modifications, peptide mass tolerance, fragment ions and specific search engine algorithms (Paulo JA, 2013). Peptide-spectrum matching (PSM) is central to protein identity inference in peptide-centric shotgun proteomics approaches. However, sequence redundancy due to the high degree of homology in protein families contributes to the indistinguishability of some proteins which match peptide sequences to multiple proteins. Ambiguous proteins, impossible to differentiate, are collapsed into a single entry or clustered as a protein group (Nesvizhskii AI and Aebersold R, 2005).

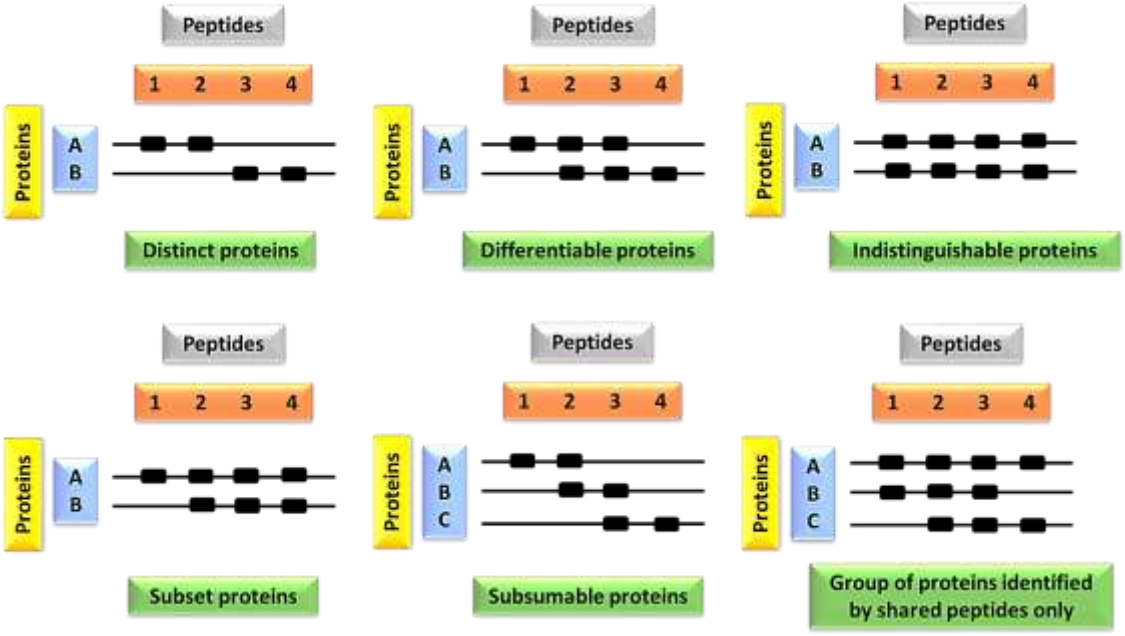

Figure 1.10: Basic peptide grouping identification scenarios that occur due to sequence homology of peptides from limited regions of the protein. (Reprinted/redrawn from open access journal with permission, (Nesvizhskii AI and Aebersold R, 2005))

Binary comparisons of sequence homology search engines (Mascot, SEQUEST, MaxQuant) show 76 - 82% peptide level overlap with a 50 - 87% protein level overlap. Variance in peptide identification can be associated with the number of spectra mapped to peptides, algorithms for peak picking, peptide sequencing, spectra passing search engine thresholds and differential assignments of peptides by search engines. Post-search processing reduces redundancy via protein grouping which assigns peptides to the protein with the highest total probability. The union of data analysed by multiple search engines, consensus methods, or the intersection of search engine results, can produce more comprehensive data sets. However, there is also a risk of increasing the false discovery rate (FDR), which in principle is directly relatable to the size of the data set (Dagda RK, Sultana T *et al.*, 2010; Paulo JA, 2013; Vaudel M, Burkhart JM *et al.*, 2011).

The union of data from four search algorithms has shown only modest improvements in sensitivity over the use of two or three search engines. Additionally, Dagda *et al.* suggest that the appropriate combination of search algorithms, not the number of search engines, maximises accuracy of peptide identification (Dagda RK, Sultana T *et al.*, 2010). Once peptides are identified in labelled proteomics, the liberated reporter ions provide quantification with protein inference influencing determination of relative abundance. Quantification limits need to be introduced to limit the range of intensities in which quantification is considered reliable. Computing protein abundance ratios firstly requires selection and filtering at the peptide level (Burkhart JM, Vaudel M *et al.*, 2011).

Combining protein quantification from different peptides to predict protein abundance estimates relative abundance by averaging the label ratios of all peptides. Unique peptides within a protein provide a direct measure of protein abundance while shared peptides require weighted averaging of abundance ratios (Burkhart JM, Vaudel M *et al.*, 2011; Nesvizhskii AI and Aebersold R, 2005). Carrillo *et al.* suggested that combining absolute peptide intensities prior to relative abundance estimates is an accurate approach to inferring protein abundance from peptides (Carrillo B, Yanofsky C *et al.*, 2010).

1.9 Hepatic Proteome

Data from the human protein atlas suggests that 61% of all human proteins are expressed in the liver with a certain proportion having elevated expression in the liver. In-depth transcript analysis of hepatic tissue shows no linear relationship between transcriptome levels and protein abundance. Lack of linear correlation demonstrates that unravelling the human liver proteome (HLP) and that of cells which are being used to mimic the liver *in vitro* is important. Draft maps of the human proteome by various groups have aided global analysis of the hepatic proteome (Kim M-S, Pinto SM *et al.*, 2014; Wilhelm M, Schlegl J *et al.*, 2014) but have, as of yet, not provided definitive insight in applications to modelling hepatocytes.

Piersma *et al.* assessed the HLP of 10 Mongolian samples using four technical approaches combining preparative two-dimensional electrophoresis, SDS-PAGE, strong cation exchange (SCX) and reverse-phase liquid chromatography (RPLC) with various analytical methods including ESI-tandem mass spectrometry (ESI-MS/MS) and MALD-TOF/TOF-MS. Implementation of these technical approaches identified 5 454 unique proteins (at 99% confidence) of which 3 013 were identified with two or more unique peptides (Piersma SR, Warmoes MO *et al.*, 2013). Additional studies comparing technical approaches are required to establish the most appropriate methods of analysing the HLP which can then be used to answer important biological questions concerning liver function and hepatotoxicity.

1.10 Research Design and Project Aim

As primary human hepatocytes are considered the "*gold standard*", any models for predicting hepatotoxicity should functionally resemble PHHs and relate to the biological questions under investigation. Implementation of technological innovation in toxico-proteomics strategies is a potential avenue for predictive preclinical safety pharmacology. Scientifically credible mathematical representations accompanying valid biological system effects are essential in pharmaceutical research and development. Importantly, adaptability to high-throughput screening and biological applicability are considered as key determinants of the feasibility of application.

The aim of the research presented in this thesis was to determine proteomic differences between pooled healthy donor primary human hepatocytes and differentially cultured human hepatocyte-derived cell lines or differentiated hepatocyte-like cells. Additionally, two cellular hepatocyte models were used to generate non-specific, proteome-wide information associated with exposure to selected known hepatotoxins to identify potential proteomic signatures of hepatotoxicity.

1.10.1 Objectives

1. Optimization of procedures for 3D culturing of HepG2 cell spheroids using hanging drops
2. Optimization of sample preparation of proteins for gel and mass spectrometry based proteomics
3. Differentiation of human induced pluripotent stem cells (iPSCs) into hepatocyte-like cells (HLCs)
4. Whole proteome analysis of the kinetics of differentiation of iPSCs to HLCs using stabled isotope labelled mass spectrometry based proteomics
    - Time course 1: iPSC (Day 1) through endoderm commitment, anterior definitive endoderm specification and hepatic progenitors to HLCs (Day 35)
    - Time course 2: Hepatocyte maturation (Day 16 to Day 40)
5. Comparison of the proteomes of terminally differentiated iPSC derived HLC monolayers and HepG2 cells as monolayers cultures as well as HepG2 three-dimensional spheroid cultures to commercially available pooled primary human hepatocytes, using stabled isotope labelled mass spectrometry based proteomics
6. Determination of the influence of selected hepatotoxic drugs, on monolayers cultures of HepG2 cells, using a multi-parametric assay model to identify sub-toxic drug concentrations
7. Determination of drug exposure specific protein expression, using stabled isotope labelled mass spectrometry based proteomics, following exposure to sub-toxic concentrations of hepatotoxic drugs in HepG2 cells cultured in both monolayers and three-dimensional spheroids

Chapter 2: Cell Culture

2.1 Materials and Methods

Chemicals, analytical reagents and cell culture medium were purchased from reputable vendors and distributors including: Merck Chemicals (Darmstadt, Germany); Sigma-Aldrich (St Louis, USA); Microsep (Waters Corporation; Milford, USA); Gibco, Lifeline Cell Technology, Thermo Fischer Scientific (Maryland, USA); Bio-Rad Laboratories, Life Science Research Division (California, USA); Thermo Fischer Scientific Inc. (Massachusetts, USA); Separations (Randburg, RSA); Qiagen (Venlo, Netherlands) and Promega (Madison, USA). Consumables were purchased from Laboratory and Scientific Equipment Company (Cape Town, RSA); AEC-Amersham (Halfway House, RSA); Corning Incorporated Life Sciences (Lowell, USA) and 3D Biomatrix (Michigan, USA). Items not purchased from the above companies were detailed where required. Unless otherwise stated, cell culture medium, enzymes and buffers were stored at 4°C. Tandem-mass tags, reconstituted cytokines and drugs were stored at -20°C or -80°C as per manufacturer's instructions.

2.1.1 Primary Human Hepatocytes

Cryopreserved pooled primary human hepatocytes (PHHs: 10 donors of varied age, gender, ethnicity, tobacco, alcohol and medication history; Lot number HUE50D) were purchased from Gibco Lifeline Cell Technology. Cryopreserved PHHs were partially thawed at 37°C in a water bath and transferred into 50 ml pre-warmed Hepatocyte Recovery Medium. Cells were centrifuged (100 *g*, 10 minutes) at room temperature, resuspended in plating medium (Williams E Medium containing 5% foetal bovine serum (FBS), 1 µM dexamethasone, 1% (v/v) penicillin-streptomycin (10 000 U/ml penicillin and 10 000 µg/ml streptomycin), 4 µg/ml human recombinant insulin, 2 mM GlutaMAX™, 15 mM HEPES (pH 7.4)) and seeded at 1 x 10^6 cells per well into a 6-well plate. Primary human hepatocytes, in suspension, were incubated at 37°C in a humidified atmosphere with 5% CO_2, for 4 hours prior to collection for proteomics experiments.

2.1.2 HepG2 Cells

Human hepatoma cells, HepG2, (85011430-1VL and CHG2-C), were obtained from the European Collection of Authenticated Cell Cultures (ECACC; United Kingdom) and Cellonex (South Africa). HepG2 cells were partially thawed at 37°C in a water bath and transferred into 25 ml pre-warmed Eagle's Minimum Essential Medium (EMEM, supplemented with 20% FBS, 1% (v/v) penicillin-streptomycin and 2 mM L-glutamine), centrifuged (100 *g*, 10 minutes) and seeded into a 25 cm^2 flask.

For passaging, cells were dissociated using trypsin-versene solution (0.25% (m/v) trypsin and 0.1% (m/v) ethylenediaminetetraacetic acid (EDTA)), for 10 minutes at 37°C. Culture medium was added, cells centrifuged (200 *g*, 5 minutes) and counted using trypan blue (0.2% (m/v)) exclusion. Cells were incubated in a humidified (5% CO_2) atmosphere at 37°C. Once sufficient stock was generated, FBS was decreased from 20% to 10% for cell maintenance and experiments.

2.1.3 Induced Pluripotent Stem Cells

2.1.3.1 Generation of Induced Pluripotent Stem Cells

Access to human induced pluripotent stem cells (iPSCs) was obtained from Professor Vallier at the Anne McLaren Laboratory for Regenerative Medicine, Wellcome Trust and Medical Research Council Stem Cell Institute and Department of Medicine, Cambridge Institute for Medical Research, University of Cambridge. An iPSC library was originally generated by Rashid *et al.* in 2010 using skin punch biopsies obtained from patients with inherent metabolic disorders (IMDs) attending Addenbrooke's Hospital, Cambridge, United Kingdom (Ethics reference number 08/H0311/201; R&D No A091485).

Briefly, fibroblasts were derived and cultured in mouse embryonic fibroblast medium (MEF medium: 450 ml DMEM, 50 ml FBS, 2 mM L-glutamine, 1% (v/v) penicillin-streptomycin). Reprogramming was achieved using four factors, OCT-4, SOX-2, c-MYC, and KLF-4 (Vectalys; Toulouse, France), under a cytomegalovirus promoter. Following derivation, hiPSCs were cultured on in-house irradiated MEF feeders in knockout serum replacement medium (KSR medium: 400 ml DMEM/F12, 100 ml knockout serum replacement (KOSR), 2 mM L-glutamine, 1% (v/v) minimum essential medium non-essential amino acids (MEM-NEAA), 0.007% (v/v) β-mercaptoethanol, 1% (v/v) penicillin-streptomycin) with 4 ng/ml fibroblast growth factor (Rashid ST, Corbineau S *et al.*, 2010).

Correction of an α1-anti-trypsin (A1AT) deficient iPSC line was achieved using a targeted biallelic gene correction of the homozygous Z mutation with a combination of zinc finger nucleases (ZFNs) and the *piggyback* transposon system to correct the point mutation (Glu342Lys; SERPINA 1). Derivation details and correction the point mutation were detailed in the publication by Yusa *et al.* and cells were grown from stocks stored under liquid nitrogen vapour (Yusa K, Rashid ST *et al.*, 2011).

2.1.3.2 Induced Pluripotent Stem Cell Thawing and Feeder Cultures

iPSCs were partially thawed at 37°C in a water bath and added to 10 ml pre-warmed KSR medium with 10 µM Y-27632 (rho-associated protein kinase inhibitor) and 4 ng/ml FGF. Colonies were centrifuged (200 *g*; 3 minutes), resuspended in KSR medium and seeded onto irradiated MEF feeder cells derived from CF1 mouse embryos (12.5 - 13.5 days post-coitum) at a density of 6.8 x 10^3 cells/cm^2.

Plates were incubated at 37°C in a humidified (5% CO_2) atmosphere with subsequent daily medium exchange. Colonies requiring maintenance on MEF feeders for more than 7 consecutive days were supplemented with 10 ng/ml activin. Colonies were passaged based on MEF age, colony size, growth patterns and differentiation. For passaging, iPSCs were dissociated with one-to-one collagenase-dispase (collagenase: 400 ml advanced DMEM/F12, 100 ml KOSR, 2 mM L-glutamine, 0.007% (v/v) β-mercaptoethanol, 0.1% (v/v) collagenase IV and 0.1% (m/v) dispase) at 37°C for up to an hour. Colonies were sedimented under gravity, washed, aspirated into smaller fragments (50 - 200 cells) and seeded repeatedly onto fresh MEFs for approximately 2 to 3 weeks from initial thawing.

2.1.3.3 Induced Pluripotent Stem Cell Maintenance

Culture plates were pre-coated with 0.1% (m/v) gelatin for 20 minutes and incubated at 37°C for 24 hours with MEF medium. iPSCs were cultured in chemically-defined medium with polyvinyl alcohol (CDM-PVA: 250 ml Iscove's Modified Dulbecco's Media, 250 ml Ham's F12 plus GlutaMAX, 1% (v/v) concentrated lipids, 7 µg/ml insulin, 15 µg/ml transferrin, 5% (m/v) PVA, 1% (v/v) penicillin/streptomycin) supplemented with activin A (10 ng/ml) and FGF (12 ng/ml). Medium was exchanged daily and colonies passaged for maintenance or differentiation at ratios of 1 to 6 and 1 to 3 respectively.

2.1.3.4 Confirmation of Induced Pluripotent Stem Cell Lineage Identity

DNA extraction was done, per manufacturer's protocol, using a GenElute™ Mammalian Genomic DNA Miniprep Kit. Harvested cells were pelleted (1 000 *g*, 2 minutes), 200 µl resuspension solution with 20 µl RNase A (2 mg/ml) was added and samples incubated at room temperature for 2 minutes. Twenty microliters proteinase K and 200 µl lysis solution C were added and samples incubated at 70°C for 10 minutes. Two hundred microliters of ethanol was added to each lysate, samples homogenized and loaded onto preconditioned binding columns. Samples were centrifuged (16 000 *g*; 1 minute), washed and spun to dryness. DNA was eluted with nuclease-free water and nucleic acid purity (260/280 and 260/230) determined using a 2 µl aliquot and a NanoDrop 1000 Spectrophotometer.

Polymerase chain reaction (PCR) was done to amplify genomic DNA containing the SERPINA 1 point mutation (536 base pairs). A ten-times master mix (7.5 µl $MgCl_2$, 7.5 µl deoxynucleotides (dNTP), 2.5 µl DNA polymerase: GoTaq (5 u/µl), 5 µl forward primer (5' TCAGAGAAAACATGGGAGGG; 0.1 µM), 5 µl reverse primer (CCCTGTTCTGAGTTGTGGT; 0.1 µM) and 100 ng extracted DNA made up to 125 µl with nuclease-free water) was made and 12.5 µl added to eight 0.2 ml PCR tubes. Thermal cycling was done using a Tetrad 2 Peltier Thermocycler at 95°C for 2 minutes, 95°C for 30 seconds and 50°C for 30 seconds for 34 cycles followed by a 5 minute cycle at 72°C and maintenance at 14°C.

DNA purification was done, as per manufacturer's instructions, using a QIAquick® PCR Purification Kit. All eight PCR samples were pooled into a single reaction mixture and 5 volumes of buffer (5 M guanidine-hydrochloride, 30% isopropanol) added. DNA was applied to a silica-based spin column (10 µg binding capacity) to recover ssDNA or dsDNA, centrifuged (16 000 *g*; 1 minute) and washed with 750 µl buffer (10 mM Tris-HCl (pH 7.5), 80% ethanol). DNA was eluted using 30 µl nuclease-free water. Amplified and purified DNA was again quantified using a NanoDrop 1000 Spectrophotometer and 4 µl was then added to the DNA loading buffer (six-times two-colour tracking dye buffer: 30% (v/v) glycerol, 1 mM EDTA, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol).

Samples were loaded onto a 1% (v/v) agarose gel with ethidium bromide and run on a Mupid-One electrophoresis unit at 100 V in Tris-acetate-EDTA buffer (TEA: 40 mM Tris-acetate, 10 mM EDTA (pH 8.2)). Gels were visualized under UV to confirm purification of genomic DNA and sent for sequencing (DNA Sequencing Facility, Department of Biochemistry; University of Cambridge, UK).

2.1.3.5 Hepatocyte-like Cell Differentiation

Hepatocyte-like cell differentiation (Figure 2.1) was based on protocols originally established by Touboul *et al.* and modified by Hannan *et al.* (Gieseck RL, Vallier L *et al.*, 2015; Hannan NR, Segeritz CP *et al.*, 2013; Touboul T, Hannan NR *et al.*, 2010).

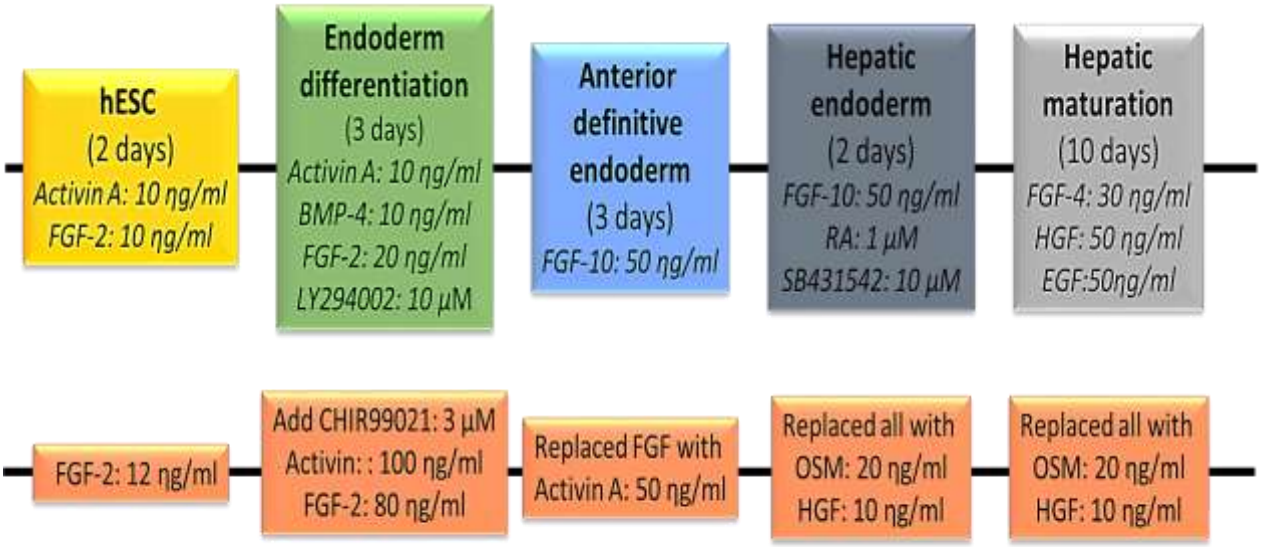

Figure 2.1: An outline of the original protocol by Touboul et al. (Touboul T, Hannan NR et al., 2010) using human embryonic stem cells (hESCs) and some of the alterations done by Hannan et al. (2013) to obtain the protocol used in this study. (Reprinted/redrawn with permission from John Wiley and Sons, license number 3961361030947, (Hannan NR, Segeritz CP *et al.*, 2013))

iPSCs, cultured in CDM-PVA with activin (10 ng/ml) and FGF (12 ng/ml) were passaged, the colonies reduced in size and seeded onto gelatine-coated plates. Plates were transferred into a tri-gas incubator (5% O_2, 5% CO_2, 90% N_2) maintained at 37°C. Differentiation, outlined in Figure 2.1 and Table 2.1, was followed until specific time course collection or terminal differentiation (Days 35 - 40).

Table 2.1: Protocol used to differentiate iPSCs to hepatocyte-like cells
(Gieseck RL, Vallier L *et al.*, 2015; Hannan NR, Segeritz CP *et al.*, 2013)

Cell specification	Day	Growth factors	Medium
hiPSC	1 - 2	Activin (10 ng/ml) FGF-2 (12 ng/ml)	CDM-PVA: Iscove's Modified Dulbecco's Medium: 250 ml Ham's F12 + GlutaMAX: 250 ml Concentrated lipids (1% v/v) Thioglycerol (20 µl) Insulin (7 µg/ml) Transferrin (15 µg/ml) Polyvinyl alcohol (PVA; 0.1% m/v) Penicillin/streptomycin (1% v/v)
Endoderm differentiation **[Endoderm commitment]**	3	Activin (100 ng/ml) FGF-2 (80 ng/ml) BMP-4 (10 ng/ml) Ly294002 (10 µM) CHIR99021 (3 µM)	
	4	Activin (100 ng/ml) FGF-2 (80 ng/ml) BMP-4 (10 ng/ml) Ly294002 (10 µM)	
Foregut endoderm **[Anterior definitive endoderm specification]**	5	Activin (100 ng/ml) FGF-2 (80 ng/ml)	RPMI 1640: B27 solution (2% v/v) Non-essential amino acid solution (NEAA; 1% v/v) Penicillin/streptomycin (1% v/v)
Hepatocyte commitment **[Hepatic progenitors]** Cells split on Day 8	6 - 10	Activin (50 ng/ml)	
Hepatocyte maturation **[Hepatocyte-like cells]** Optional cell split on Day 25	11 - 40	Oncostatin M (20 ng/ml) HGF (10 ng/ml)	Hepatozyme: Glutamine (2 mM) Penicillin/streptomycin (1% v/v) Concentrated lipids (2% v/v) Transferrin (30 µg/ml) Insulin (15 µg/ml) NEAA (2% v/v) Penicillin/streptomycin (1% v/v)

Following foregut endoderm induction (Day 8), cells were dissociated to a single cell suspension using enzyme-free cell dissociation buffer. Cells were centrifuged (200 *g*; 3 minutes), counted and reseeded onto gelatine-coated plates at a density to ensure confluence (approximately 2 x 10^5 cells/cm^2). Cells were then maintained for the duration of differentiation unless a high proportion of undifferentiated cells were observed where cells could be split on Day 25 and reseeded at confluence.

2.1.3.6 Molecular Characterisation of Hepatocyte-like Cell Differentiation

2.1.3.6.1 Primer Design

Primers were designed by Dr C Segeritz using reference gene sequences from the Ensemble Genome Browser (*http://www.ensembl.org/*) and the National Centre for Biotechnology Information (NCBI) primer designing tool (*http://www.ncbi.nlm.nih.gov/tools/primer-blast/*). Sequences were 15 - 30 nucleotides in length, flanked by known introns and contained 40 - 60% uniformly distributed GC with an annealing temperature near 60°C. Self-complimentary motifs, complementary forward-reverse pairs and direct nucleotide repeats were avoided. Specificity was assessed using the Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi). Oligopeptides were obtained at a 0.025 µmole synthesis scale (Sigma Aldrich; UK). Primer pair specificity was confirmed by single peak amplicons upon thermal dissociation (Segeritz C, 2015).

2.1.3.6.2 RNA Quantitation of Molecular Markers

RNA was extracted using 100 - 350 µl lysis solution for RNA extraction (10 mM EDTA, 0.5% (m/v) SDS, 10 mM Tris (pH 7.5) in RNase-free water) and lysates stored at -80°C. qPCR for A1AT, albumin (ALB), α-fetoprotein (AFP), CYP1A2, CYP3A4, CYP3A5, CYP3A7, CYP7A1 and hepatocyte nuclear factor 4α (HNF4α) was done using ubiquitin C (UBC) as a housekeeping gene. Molecular marker quantitation was done in biological replicates of HLC differentiations and isolated PHHs compared to iPSCs.

RNA extracts were thawed and processed, as per manufacturer's instructions, using a GenElute Mammalian Total RNA Miniprep Kit. Equal volumes of 70% ethanol were added to each RNA lysate, samples homogenized and loaded onto binding columns. Samples were centrifuged (16 000 *g*; 1 minute) and columns washed with 500 µl wash solution. On-column DNA digestion was done for 15 minutes at room temperature using 10 µl deoxyribonuclease-I diluted in 70 µl DNase digest buffer. Samples were washed, spun to dryness (16 000 *g*; 3 minutes) and eluted with 50 µl nuclease-free water. Nucleic acid purity (260/280 and 260/230) was determined using a NanoDrop 1000 Spectrophotometer.

Real-time qPCR was performed as two independent reactions to firstly generate global cDNA using non-specific primers and then amplify with gene-specific primers. For RT1, a ten-times master mix (5 µl random primer, 10 µl dNTPs) was made and 1.5 µl per reaction added to 500 ng RNA made up to 11.875 µl with dH_2O. This was run for 5 minutes at 65 °C using a Tetrad 2 Peltier Thermocycler.

A ten-times master mix (40 µl of five-times first strand buffer containing 250 mM Tris-HCl (pH 8.3), 15 mM $MgCl_2$, 375 mM KCl, 20 µl of 0.1M DTT, 5 µl recombinant RNA inhibitor, 1.25 µl superscript II reverse transcriptase) was then made and 6.625 µl added to the RT1 reaction mix. Thermal cycling was done at 25°C for 10 minutes, 42°C for 50 minutes, 70°C for 15 minutes and maintained at 4°C. For qPCR, 10 µl of ten-times master mix (75 µl Bioline SensiMix SYBR, 13 µl dH_2O, 0.6 µl forward primer, 0.6 µl reverse primer) was added to 5 µl diluted cDNA in duplicate per primer pair. Incorporation of SYBR Green-I was measured using an Mx3000P qPCR System. The thermal profile consisted of denaturing at 95°C for 10 minutes followed by 40 cycles of 30 seconds at 95°C, 1 minute at 55°C and 1 minute at 72°C. Lastly, one cycle of 1 minute at 95°C, 30 seconds at 55°C and 30 seconds at 72°C. Initial copy numbers were determined based on the threshold cycle (Ct or crossing point) value which is inversely proportional to the log of the initial copy number.

Data normalisation was done using the Delta Delta Ct (Livak method):

$$\Delta\Delta CT = [\mathrm{Ct}(\mathrm{target}, \mathrm{untreated}) - \mathrm{Ct}(\mathrm{reference}, \mathrm{untreated})] - [\mathrm{Ct}(\mathrm{target}, \mathrm{treated}) - (\mathrm{Ct}(\mathrm{reference}, \mathrm{treated})]$$

Calculating the ratio of target genes was done by calculating the $2^{\Delta\Delta CT}$. The ΔCt (target – reference) of targets was raised to the power two and data for HLCs and PHHs normalized to iPSC gene expression.

2.1.3.7 Cytochrome P450 Assay

Cytochrome P450 activity is a key indicator of HLC maturity and functionality. CYP3A4, was assessed over a time-course using a non-lytic Luciferin-isopropyl acetal (Luciferin-IPA) P450-Glo™ CYP3A4 assay. Luciferin-IPA is a pro-luciferin converted by CYP3A4 to luciferin which interacts with luciferase enzymes in buffer detection reagents. Lyophilized luciferin detection reagent was solubilized in esterase-containing reconstitution buffer and stored at -20°C.

Luciferin-IPA substrate (3 µM in dimethyl sulfoxide), was diluted (1:1000) in hepatozyme containing 20 ng/ml oncostatin M and 10 ng/ml HGF. Monolayers were incubated at 37°C for 30 minutes with 500 µl of luciferin-IPA substrate-containing medium. Fifty microliters of medium and detection reagent was added, in triplicate, to opaque 96-well plates. Following gentle agitation and incubation at room temperature for 20 minutes, luminescence was detected using a GloMax® Microplate Luminometer. Relative luminescence units (RLU) were blank-adjusted, normalized to total protein (Chapter 3.1.3: Protein Quantitation of Spheroids, Bicinchoninic acid assay) and expressed as a fold-change.

2.2 Results and Discussion

PHHs and HepG2 cells represent a diverse background of commonly used cellular sources with applications in safety pharmacology. PHH and HepG2 cells were cultured using standard protocols which required no additional characterisation. Therefore, here the characterisation of HLC differentiation is the primary discussion focus. Gene targeting technology for simultaneous biallelic correction via ZFNs and the *piggyback* transposon system combination was used for wild-type correction of iPSCs as it prevents residual non-human sequence contamination of the host genome. Although corrected iPSCs carry 29 mutations in protein-coding exons, 22 were splice site mutations or non-synonymous. Wild-type corrected iPSCs were successfully differentiated to HLCs suggesting limited biological consequences induced by the genomic abnormalities (Yusa K, Rashid ST *et al.*, 2011). These generated iPSC were successfully thawed and cultured for 2 - 3 weeks on MEF feeders prior (Figure 2.2A - C) to seeding stabilized colonies onto gelatin-coated plates (Figure 2.2D - E).

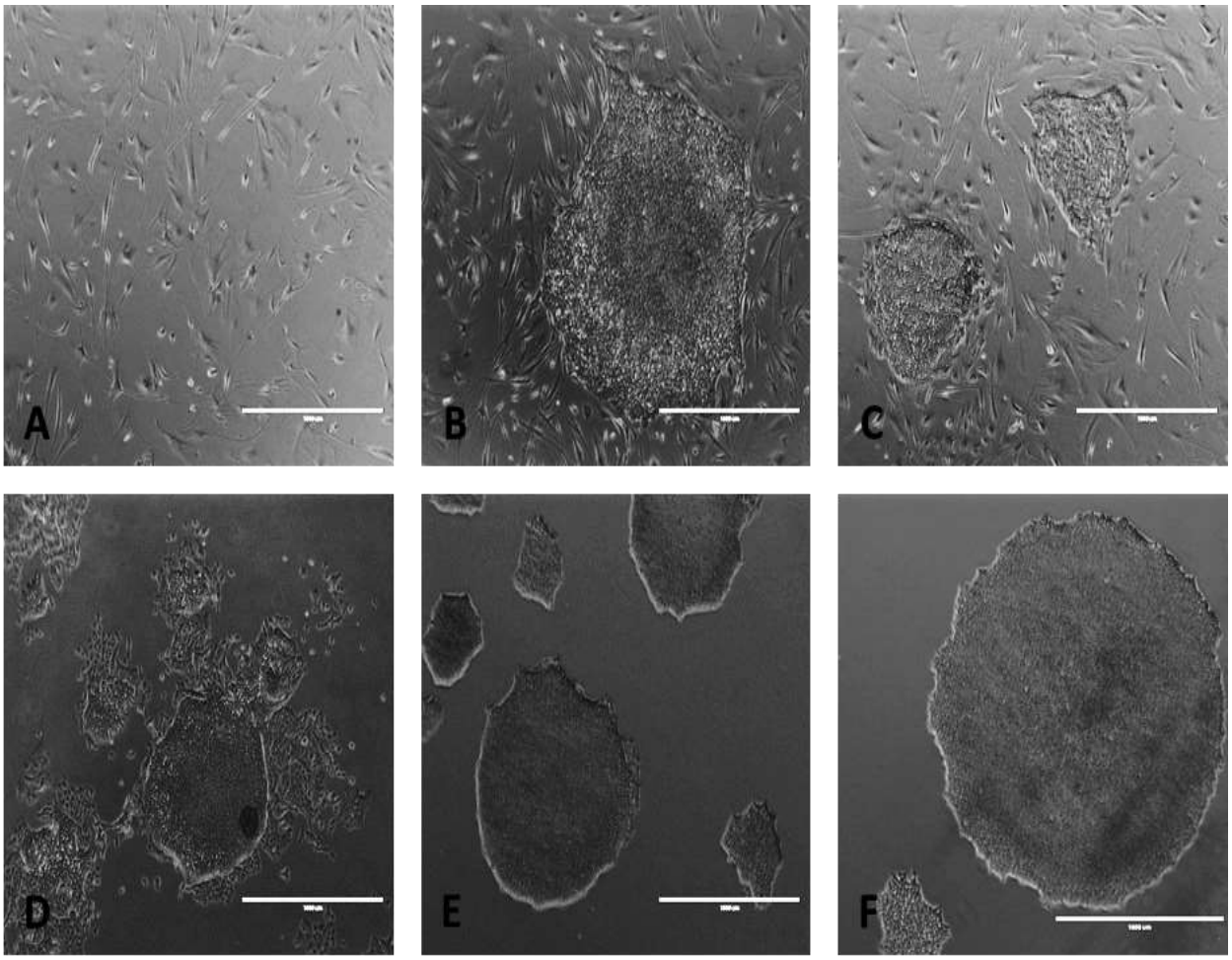

Figure 2.2: Phase contrast images (EVOS FL Cell Imaging System, scale bar: 1000 µm) of A) Mouse embryonic fibroblast (MEF) feeder cells, B - C) iPSC colonies growing on MEF feeder cells, D) iPSC colonies at the first passage on feeder-free gelatin-coated plates illustrating a high degree of spontaneous differentiation, E - F) Stabilised iPSC colonies used for differentiation.

Due to the use of iPSCs from an already established iPSC library, the genome of stable colonies was sequenced to confirm iPSC lineage identity as to ensure the use of a wild-type corrected cell line. Initial amplification of the 536 base pair SERPINA 1 gene to be sent for sequencing produced multiple bands (Figure 2.3A) suggesting inefficient primer binding efficiency. Enhancing primer binding was done using a different DNA polymerase at a higher concentration which produced a single amplified band (Figure 2.3B) which was sent for sequencing (DNA Sequencing Facility, Department of Biochemistry; University of Cambridge, UK).

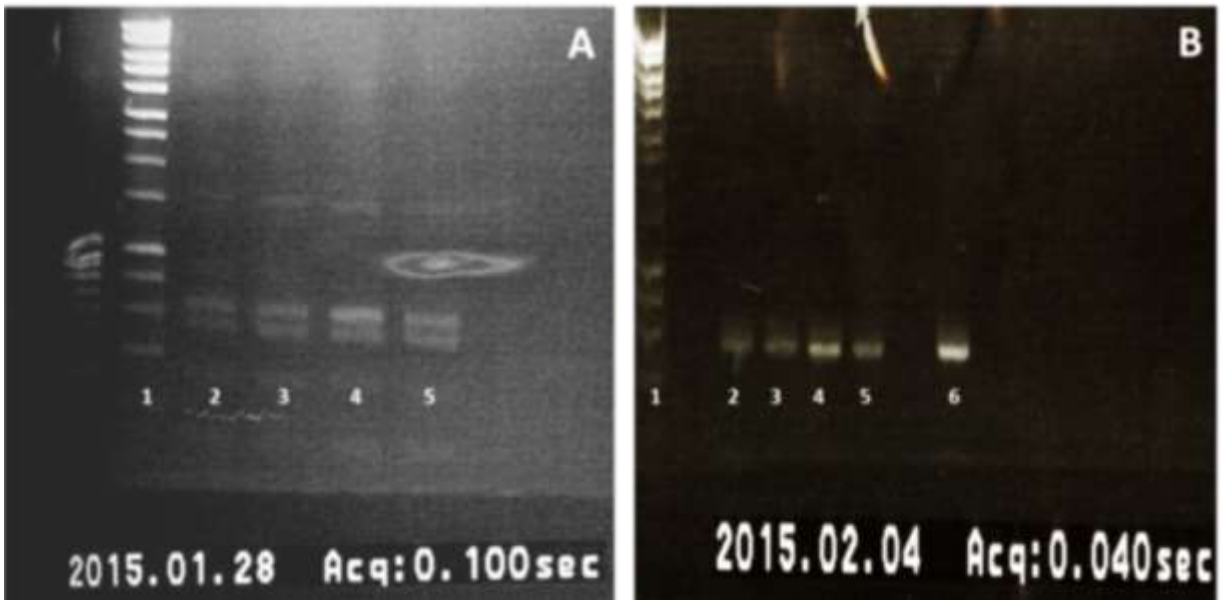

Figure 2.3: Agarose gels of various iPSC lines. A) Inefficient gene amplification with multiple bands per sample. Lanes were loaded with 1) DNA ladder (1kB: 10 000 to 250), 2 - 4) Mutant iPSC lines, 5) Corrected iPSC lines and B) Successful SERPINA 1 amplification with a single bands at approximately 500 kb. Lanes were loaded with 1) DNA ladder, 2 - 4) Mutant iPSC lines, 5) Corrected iPSC lines, 6) Pooled sample.

Sequence data for IMD mutant and wild-type corrected iPSC lines was obtained, with 536 bases analysed in 6 462 scans. The number of base pairs detected suggests that the SERPINA 1 had been amplified without additional non-specific amplification. The A1AT point mutation was identified as the first base pair in the codon prior to the AAAGGG sequence coding lysine. Guanine (GAG) here encodes for glutamine which is the wild-type, whereas a point mutation to adenine (AAG) encodes for lysine. Sequencing identified the wild-type correction in the iPSC line being cultured which confirmed use for further experimentation (Figure 2.4).

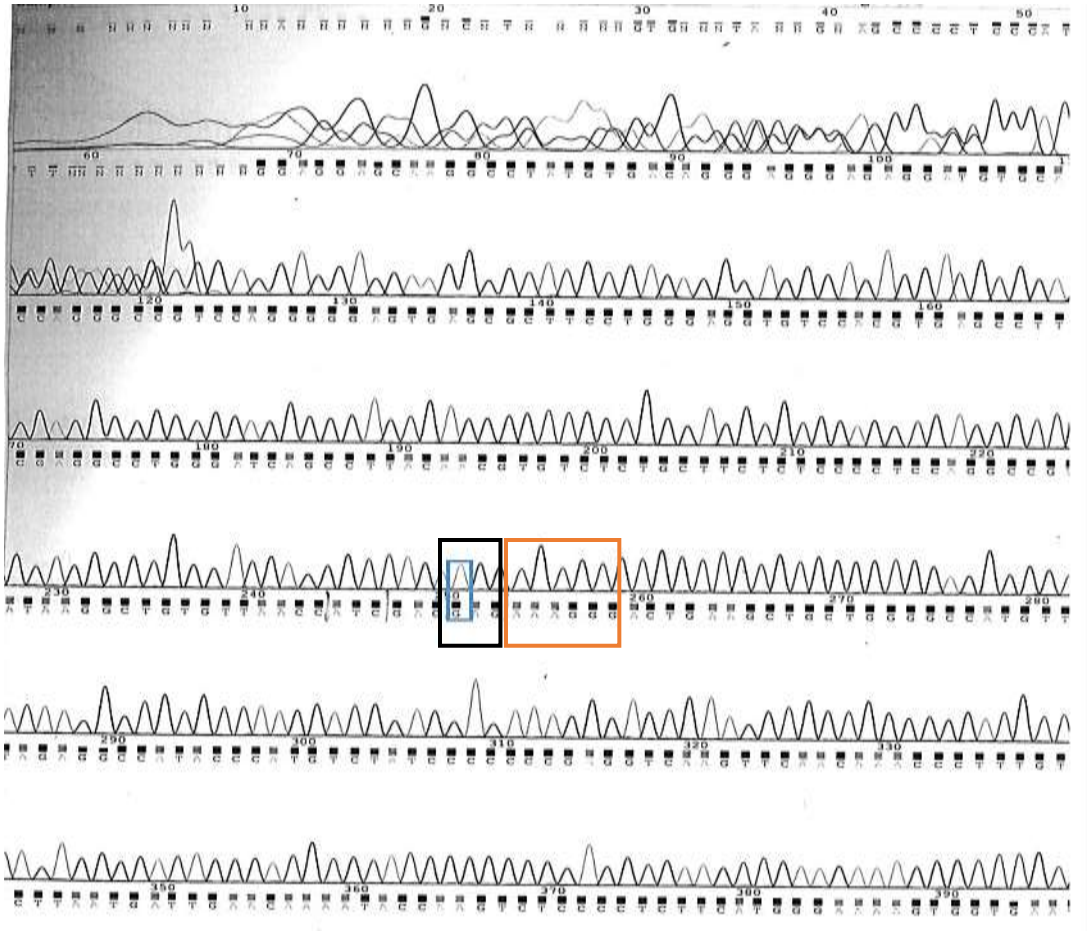

Figure 2.4: Sequence data of the PCR amplified SERPINA 1 gene corrected iPSCs. The AAAGGG sequence (orange box) is used as an identifier for the codon of interest (black box). Guanine (indicated by the blue box) is the sequence for the wild type, which is substituted to adenine in the mutant lines resulting in a lysine in place of wild-type glutamine.

Despite potential genetic instability occurring during cell reprogramming, differentiation to HLCs provides an unlimited source of material for preclinical models. Lessons from developmental studies broadly infer suitable markers and signalling factors for confirmation of differentiation. Compartmentalization from one developmental stage to another is not observed during differentiation due to cross-regulation and auto-regulation of transcription factors (Gieseck RL, Vallier L *et al.*, 2015; Lemaigre F, 2010). Multistep protocols initially differentiate iPSCs into definitive endoderm which is the common progenitor for the liver, lungs, pancreas, gut and thyroid. Development and patterning of regionalized endoderm into anterior definitive endoderm initially limits lineage capabilities. Differentiation of foregut endoderm into hepatoblasts and propagation of bipotent hepatoblasts into hepatocytes lineages demonstrate distinct temporal gene expression profiles (Gieseck RL, Vallier L *et al.*, 2015; Lemaigre F, 2010).

The natural highly co-ordinated developmental pattern provides a broad guideline for hepatocyte differentiation with numerous options for manipulation (Hay DC, Zhao D *et al.*, 2008; Moore RN, Dasgupta A *et al.*, 2008; Sharma NS, Wallenstein EJ *et al.*, 2009; Si-Tayeb K, Noto FK *et al.*, 2010; Song Z, Cai J *et al.*, 2009; Takayama K, Inamura M *et al.*, 2012; Touboul T, Hannan NR *et al.*, 2010). Commonality in induction of differentiation protocols is evident while specification, matrices and culture approach often differ (Figure 2.5). It is unclear as to which protocol results in terminally differentiated, mature HLCs with "benchmark" comparisons of morphological, phenotypic and functional analyses being required (Schwartz R, Fleming H *et al.*, 2014). Despite this, cells derived via these protocols generally mimic human hepatocytes with upregulated proteins more closely resembling the liver as opposed to any other organ or tissue (Baxter M, Withey S *et al.*, 2015).

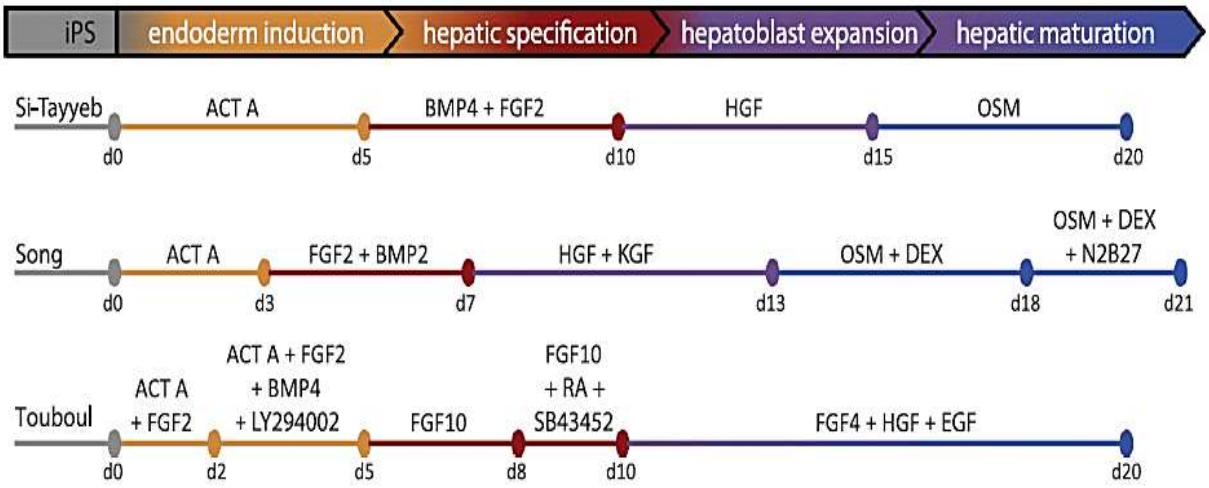

Figure 2.5: Three stepwise protocols for generation of hepatocyte-like cells from iPSCs. Schematic outlining the differentiation kinetics and growth factors utilized in three different but commonly used protocols (Si-Tayeb K, Noto FK et al., 2010; Song Z, Cai J et al., 2009; Touboul T, Hannan NR et al., 2010). (Reprinted with permission from Elsevier, license number 3961370418748, (Schwartz R, Fleming H *et al.*, 2014))

Expression of cross-regulating core transcription factors and hepatic phenotypes normally characterised at different stages of the differentiation process are reasonably well defined (Figure 2.6). However, characterisation should be far more multidimensional. Schwartz *et al.* recommended the following essential assessments for HLCs: 1) morphologically cells should be cuboidal or polygonal with enlarged nuclei and high cytoplasmic ratio, abundant organelles, presence of microvilli, vesicles and junctional complexes, 2) gene and protein expression to ensure correct lineage specification and 3) functional characterisations which spans over 500 functional classes including bile production, protein synthesis, detoxification and energy metabolism (Schwartz R, Fleming H *et al.*, 2014).

The protocol used in this study, as well as prominent markers used to assess differentiation are outlined in Figure 2.6. Despite implementation of a well-defined and characterised protocol, genetic characterisation of HLCs was done to accompany the morphological attributes observed throughout differentiation to ensure a satisfactory outcome in accordance with that previously demonstrated using this differentiation protocol. Morphologically, differentiation was observed from iPSCs through endoderm differentiation to mature HLCs and this sequence is illustrated in Figure 2.7. Additionally, HLCs morphology was most comparable to primary human hepatocytes at approximately Day 35 despite being slightly smaller and lacking the same degree of nucleation (Figure 2.8).

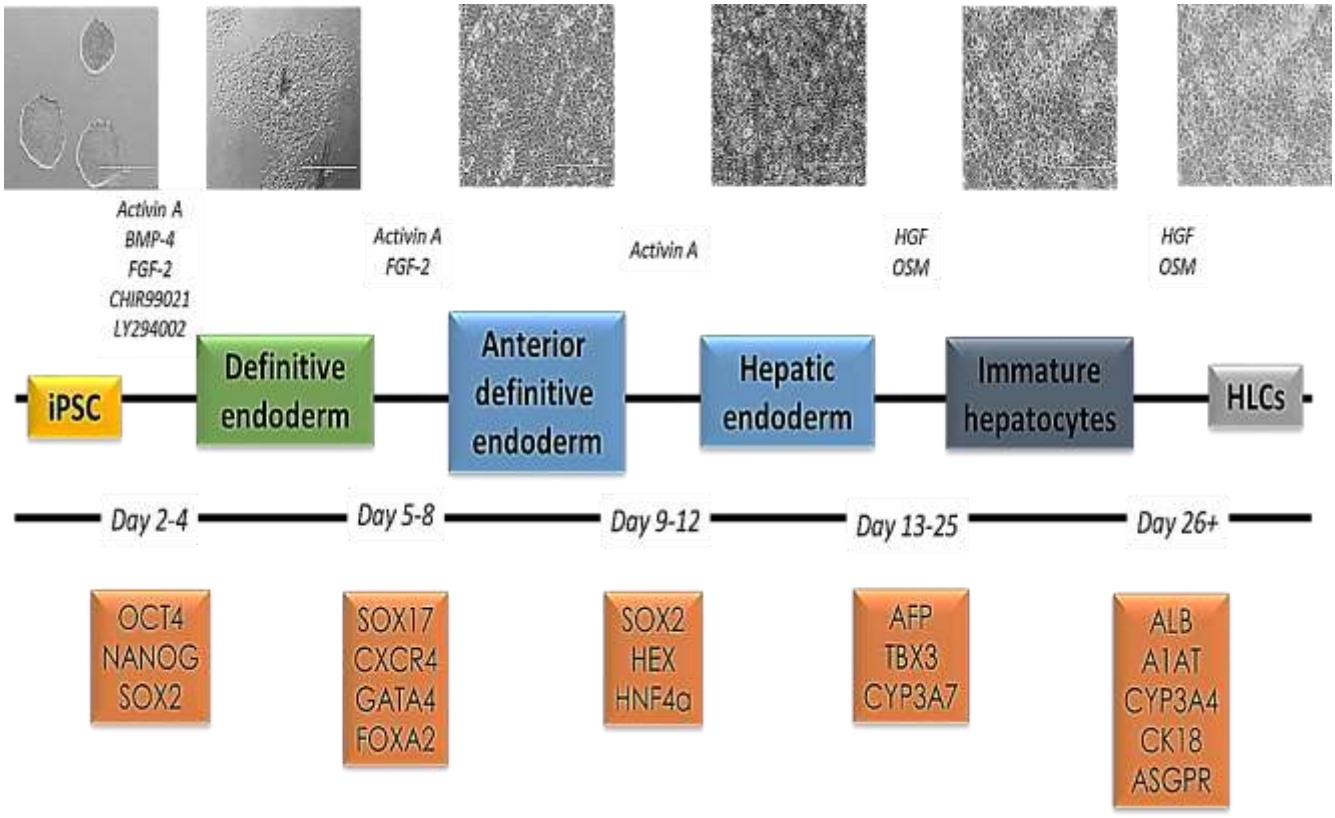

Figure 2.6: Overview of the differentiation protocol, accompanied by phase contrast images (EVOS FL Cell Imaging System, scale bar: 1000 µm). Human iPSCs undergo stepwise differentiation into hepatocyte-like cells. Pluripotent stem cells, marked by octamer-binding transcription factor 4 (OCT4), NANOG homeobox transcription factor (NANOG) and SRY-related high-mobility-group box 2 (SOX2) first differentiate into definitive endoderm expressing SOX17, CXC chemokine receptor 4 (CXCR4), GATA binding protein-4 (GATA4), and forkhead box A2 (FOXA2). Definitive endoderm is committed to anterior definitive endoderm (ADE), marked by upregulation of SOX2, hematopoietically expressed homeobox protein (HEX) and hepatocyte nuclear factor 4α (HNF4α). Patterning into hepatic endoderm precedes production of immature hepatocytes with a foetal identity, are marked by α-fetoprotein (AFP), T-box transcription factor-3 (TBX3) and CYP3A7. Maturation into hepatocyte-like cells is characterised by markers, including albumin (ALB), α1-antitrypsin (A1AT), CYP3A4, cytokeratin 18 (CK18), and asialoglycoprotein receptor (ASGPR). (Reprinted/redrawn from open access journal with permission, (Gieseck RL, Vallier L *et al.*, 2015))

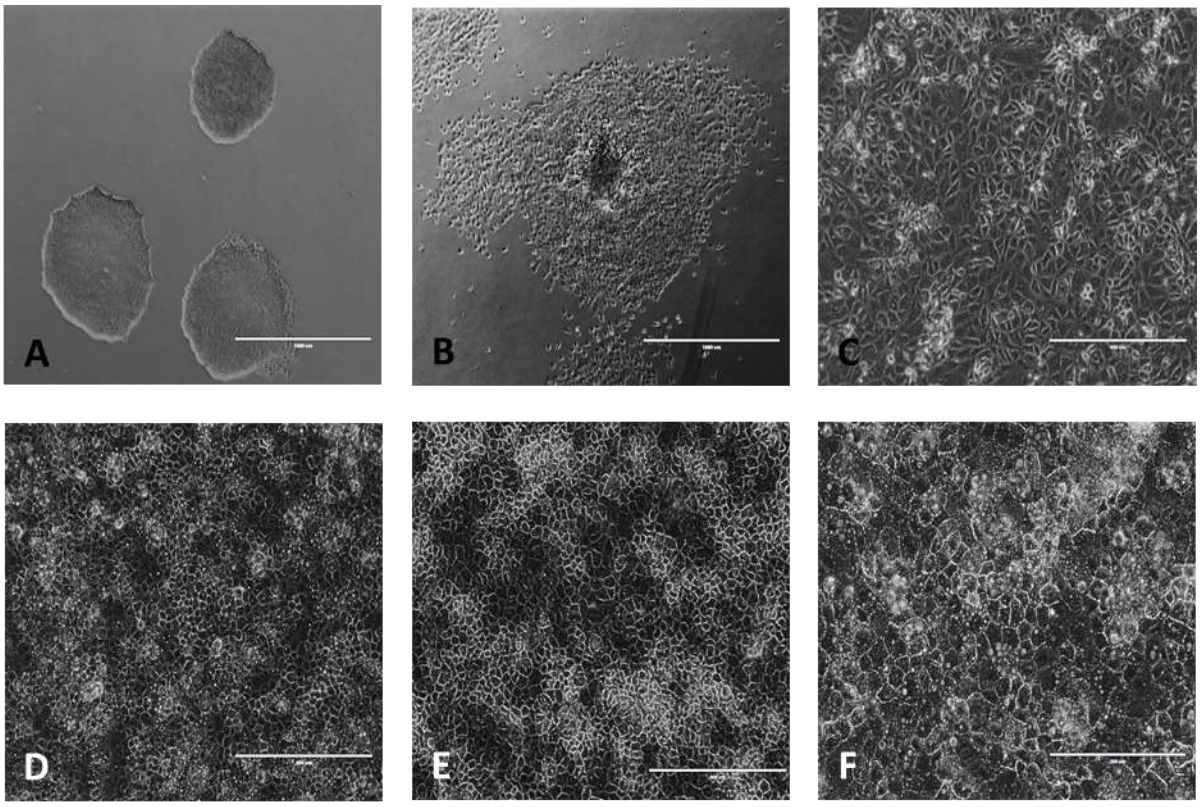

Figure 2.7: Phase contrast images (EVOS FL Cell Imaging System) through the differentiation time course. A) Human iPSCs (not at differentiation density) which have a high nuclear to cytoplasmic ratio, B) Foregut endoderm at Day 6 where cells enlarge and migrate outward, C) Hepatic progenitors at Day 9 developing a fibroblastic morphology, D) Hepatocyte-like cell maturation at Day 16 where a polygonal morphology and distinct cell boundaries are evident, E) HLC maturation at Day 25 and F) Mature HLC at Day 35 with prominent nucleoli and potential binucleation. Scale bar: 1000 µm.

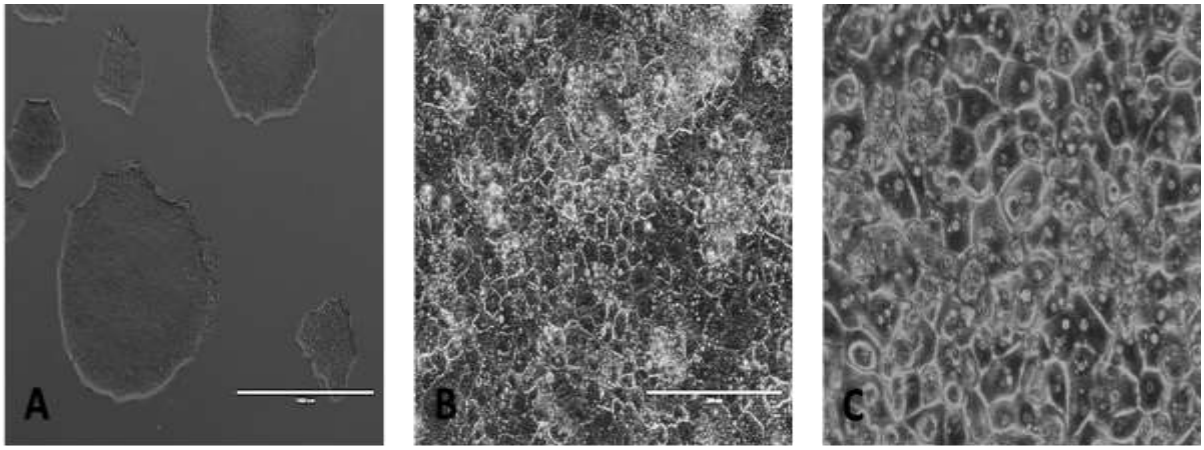

Figure 2.8: Phase contrast images (EVOS FL Cell Imaging System) of A) Human iPSCs (not at differentiation density), B) Hepatocyte-like cells differentiated to Day 35 and C) Primary human hepatocytes which are cuboidal in shape, display polarization and have bile canaliculi between adjacent cell membranes. Scale bar: 1000 µm.

Hepatoblasts downregulate α-fetoprotein and CYP3A7 expression, which are considered characteristics typical of foetal phenotypes. Contrastingly, expression of albumin, A1AT and CYP3A4 are associated with mature hepatocytes (Figure 2.9). Characteristics of foetal and mature phenotypes have been observed with the protocol used in this study. However, detailed characterisation demonstrates limited or absent expression of essential drug transporters and deficient polarization in monolayers resulting in an overall lack of mature phenotype (Gieseck RL, Vallier L *et al.*, 2015).

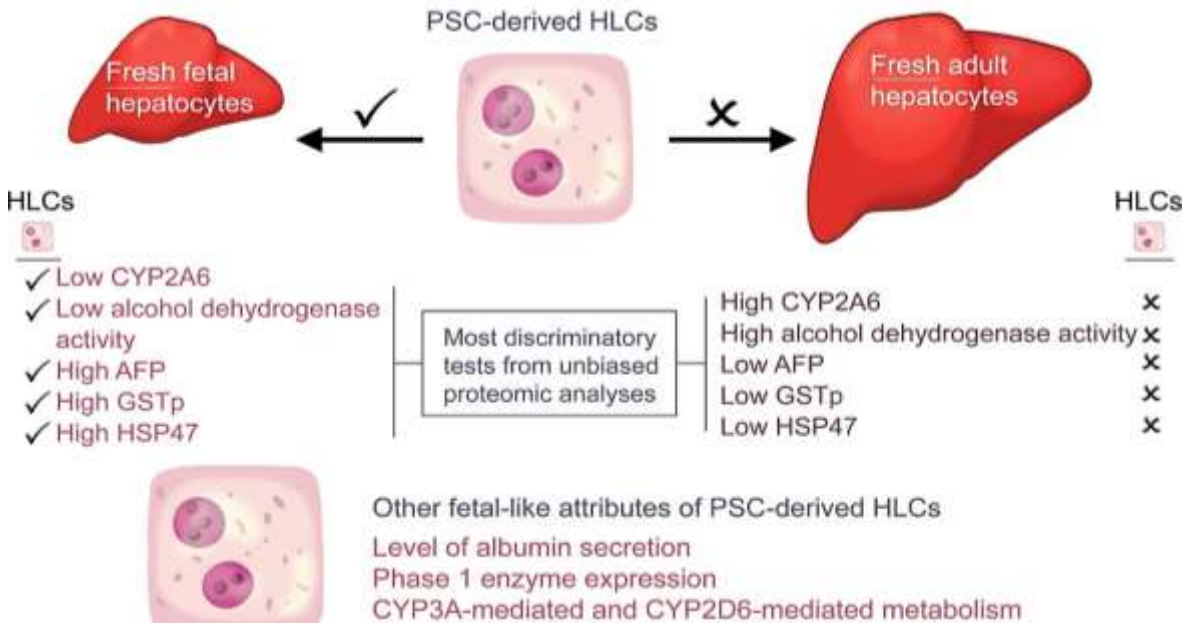

Figure 2.9: How mature are pluripotent stem cell-derived HLCs differentiated using soluble factors? A comparison between the key factors distinguishing HLCs from fresh foetal and adult hepatocytes. The resulting outcomes suggesting a strong correlation with foetal hepatocytes. (Reprinted from open access journal with permission, (Baxter M, Withey S *et al.*, 2015)).

Baxter *et al.* suggested that limited maturity using various protocols raises two pertinent questions. Firstly, are cells aberrant because of inadequate specificity in following human liver development? Or considering correct lineage specificity, are HLCs trapped in a foetal-like state? Assessing the similarity of HLCs and adult hepatocytes is challenging as PHHs dedifferentiate rapidly and show high variability. Additionally, HLC maturity exists on a scale from undifferentiated foetal hepatocytes to differentiated primary adult hepatocytes and should therefore be benchmarked against both phenotypes. However, access to foetal hepatocytes for experimental purposes remains limited. This could result in over-interpretation of HLC phenotypes (Baxter, Withey et al. 2015). Despite this, proteome-wide analysis could highlight both the similarities and limitations of HLCs compared to other models. Additional characterisation was done to ensure correct lineage specification prior to proteomic assessments.

Inductive cues for differentiation promote sequential acquisition of the nuclear transcription factor HNF4α in hepatic progenitors, AFP in hepatoblasts and albumin in hepatocytes. HNF4α, which appears only during hepatic lineage specification, controls hepatic gene expression and is crucial in maintaining transcription factor networks essential to hepatocyte function. As a master regulator, HNF4α induces other dynamically altered genes in hepatocyte formation including FOXA2, TBX3, HHEX, HNF4α, GATA4 and GATA6 (Figure 2.10)(DeLaForest A, Nagaoka M *et al.*, 2011). For characterisation, mRNA quantification of hepatic markers albumin, A1AT, AFP and HNF4α as well as metabolizing enzymes, CYP1A2, CYP3A4, CYP3A5, CYP3A7 and CYP7A1 was compared to PHHs and iPSCs.

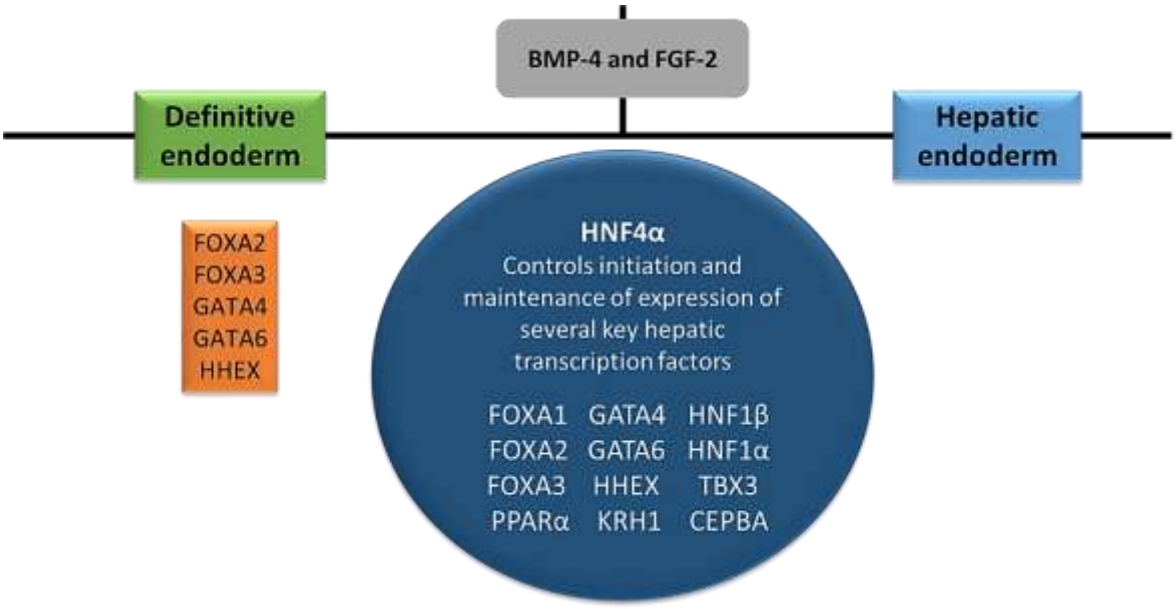

Figure 2.10: Schematic of the role of HNF4α during the differentiation of human ES cells into hepatocyte-like cells. Definitive endoderm generated from human ES cells expresses key hepatic transcription factors, including FOXA2, FOXA3, GATA4, GATA6 and HHEX that, based on studies in the mouse, probably regulate the ability of the endoderm to adopt a hepatic destiny. In response to inductive cues, including FGFs and BMPs, the receptive endoderm expresses HNF4α within the nascent hepatic progenitor cells where it is responsible for the initiation and maintenance of expression of several hepatic transcription factors, including HNF1β, that control formation of hepatoblasts and their differentiation towards functional hepatocytes. (Reprinted/redrawn from open access journal with permission, (DeLaForest A, Nagaoka M *et al.*, 2011))

Limited replicates of mRNA quantification were conducted with the use of primary human hepatocytes from only three isolations and three biological replicates of HLC differentiation. HLC expression of A1AT was approximately half that of PHHs, albumin expression was roughly one-sixtieth that of PHHs, and expression of HNF4α was one-third that of PHHs (Figure 2.11). Basal transcript level of AFP was low in PHH but highly expressed in HLCs. Despite HNF4α gene expression being low, this coupled to detection of albumin and AFP suggests transition to hepatic progenitors.

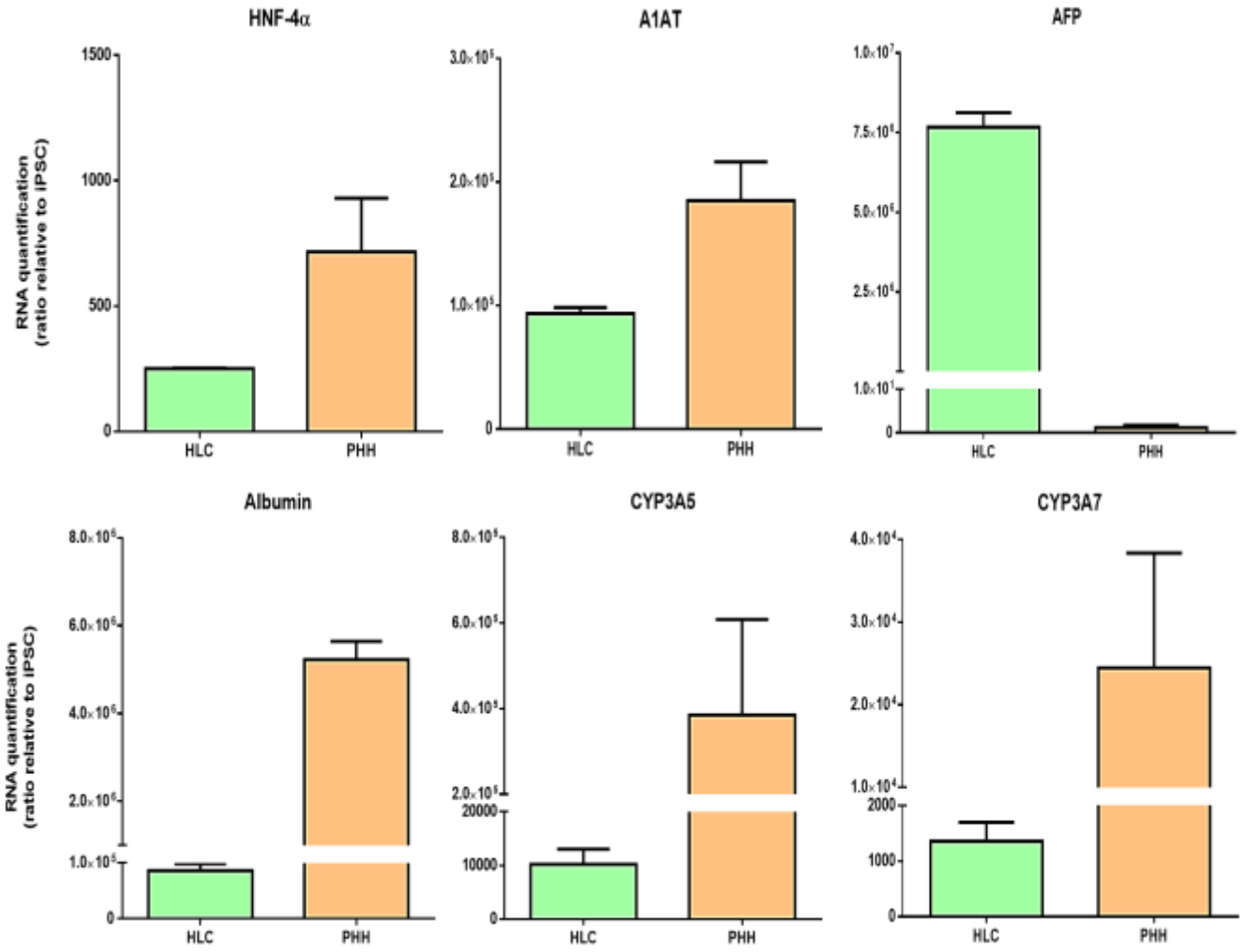

Figure 2.11: qPCR data for common hepatic markers hepatocyte nuclear factor 4α *(HNF4α), α1-antitrypsin (A1AT), α-fetoprotein (AFP), albumin, CYP3A5 and CYP3A7 comparing HLCs from three differentiations to isolated PHHs. Data was expressed as a ratio relative to iPSCs with differences all being statistically significant with* $p < 0.05$.

Identification of more authenticated regulators of hepatocyte differentiation will aid in tailoring protocols towards maturation. Furthermore, correlation of gene expression to proteomic profiles may prove insightful into the advantages and shortcomings of HLCs. Slight increases from basal iPSC expression of CYP3A5, CYP3A7 and CYP7A1 were evident but this did not rival expression in PHHs. Limited expression of these enzymes at the mRNA level are of concern as six CYP450 enzyme isoforms are thought to collectively account for approximately 95% of CYP450-mediated reactions. Notably CYP3A4 accounts for up to 36% of xenobiotic metabolism in normal liver tissue (Dambach DM, Andrews BA *et al.*, 2005; Zimmerman H, 1999) which makes, CYP3A4 activity a metabolic qualification marker of HLC maturity.

HLC monolayers have been maintained for up to 55 days (Gieseck RL, Hannan NR *et al.*, 2014), however, the typical endpoint is between 35 and 40 days. A time-course of CYP3A4 activity was conducted every alternate day for 12 days from Day 25, to define the collection endpoint for this study. Normalized CYP3A4 activity increased 3-fold from Day 31 to 35 of hepatic maturation. An initial decline was observed at Day 37. Despite increases in CYP3A4 activity, the metabolic relationship to PHHs is still considered weak. However, based on these CYP activity results, collection of differentiated samples cultured as monolayers was done at Day 35 for further assays including proteomic investigation.

A noteworthy limitation of hepatic comparisons is the rapid degree of dedifferentiation of PHHs as well as the remarkable variability in maturation between differentiations. HLC phenotypes are reported to be shifted closer towards PHHs using Real Architecture for 3D Tissues (RAFT) systems which make use of a concentrated cell-seeded collagen hydrogel (Figure 2.12). Maintenance of cell-cell junctions by seeding small epithelial clumps enhances hepatic maturation enabling culturing for 75 days compared to 55 days for 2D counterparts (Gieseck RL, Hannan NR *et al.*, 2014).

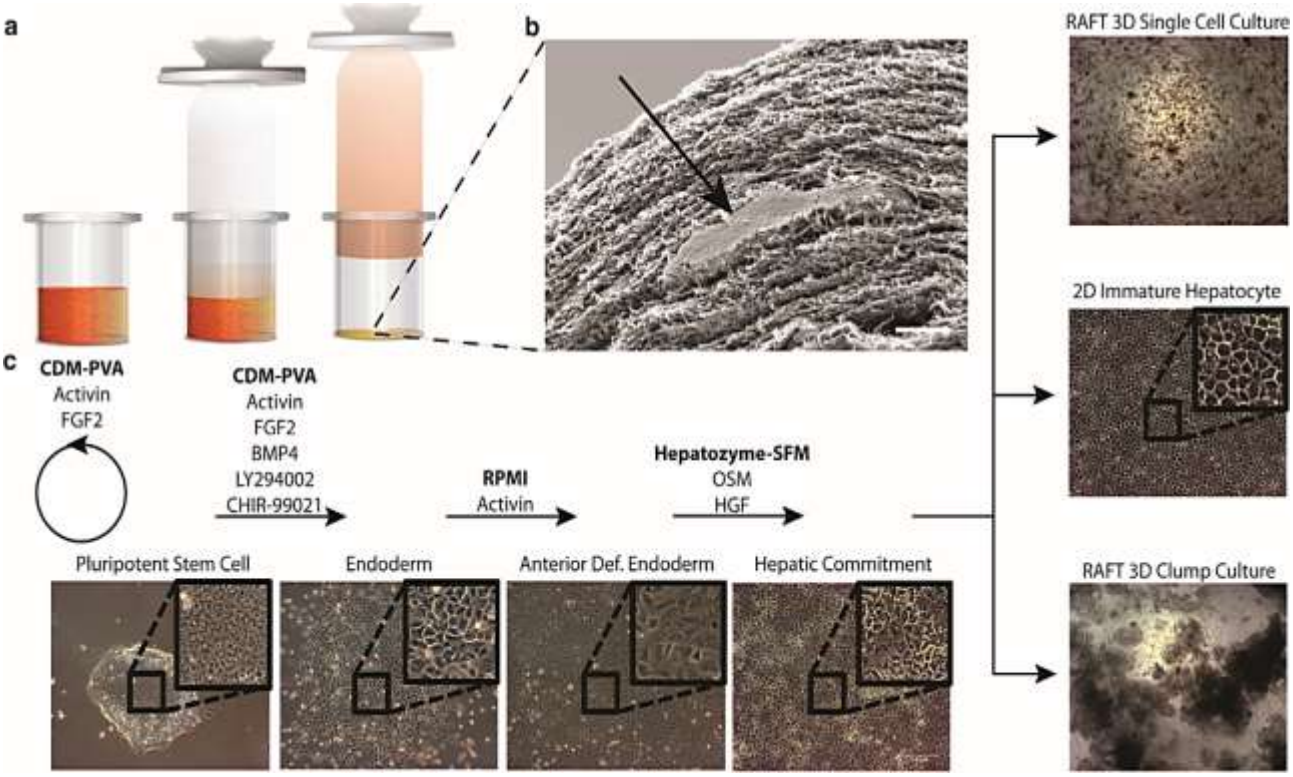

Figure 2.12: Method to differentiate iPSC-HLCs in the RAFT system. a) Schematic of the RAFT process used in the maturation of IPSC-HLCs, b) Scanning electron micrograph of 3D clump culture and c) Outline of the experiment used to probe the effects of 3D culture conditions on the maturation of iPSC-HLCs. (Reprinted from open access journal with permission, (Gieseck RL, Hannan NR *et al.*, 2014))

Three dimensional models appear to be important in promoting hepatic phenotypes. However, cellular extraction from collagen matrices is difficult and would not yield suitable material for proteomics. While 3D cultures have been performed using HLCs generated using this differentiation protocol, the influence of culture format on maturation cannot be ignored. Due to the different morphology and format of 3D hepatocyte-like cell cultures, it was not considered feasible to extrapolate characterisation and maturation data between 2D and 3D culturing techniques. Therefore, in this study only proteomic comparisons to monolayer cultures of HLCs, where adequate supporting data was available, were conducted.

2.3 Conclusion

PHHs, HepG2 cells and HLCs represent diverse backgrounds of cellular sources for potential application in assessing safety pharmacology features associated with the liver. Cryopreserved pooled PHH were purchased due to limited access to freshly isolated liver tissue. Furthermore, commercially available HepG2 cells are reputed to be phenotypically suppressed yet easy to obtain and culture. Biological insights from HLCs in drug discovery is based on the ability to mimic PHHs. Assessing the proteome during differentiation in addition to comparing to other potential cell types is essential in determining whether HLCs can be developed into a robust preclinical hepatotoxicity model. The differentiation protocol conducted here is well-established and tailored to promote hepatic maturity.

HLC differentiation from various laboratories still produce more pronounced similarities to foetal hepatocytes (Baxter M, Withey S *et al.*, 2015; DeLaForest A, Nagaoka M *et al.*, 2011). Regardless of the limitations in the ability to generate mature hepatocytes, these cells are considered one of the best sources available for regenerative medicine and drug discovery endeavours. One of the major advantages of the differentiation protocol is the avoidance of feeder cultures or embryoid bodies (during differentiation) which are thought to decrease the reproducibility and synchronicity of differentiation. Molecular characterisation supports the differentiation of hepatic phenotypes with results comparable to others using the same protocol. Competence in differentiation from iPSCs was achieved and these cells were used to assess the kinetic changes in differentiating HLCs and comparative proteomic studies.

2.4 Appendix: Cell Culture

Table 2.2: Medium and other constituents used for cell culture

Component	Volume	Final concentration	Company	Catalogue number
HepG2 hepatocellular carcinoma cells				
Eagle's Minimum Essential Medium (EMEM)				
EMEM (from powder)	9.3 g/l	-	Sigma Aldrich	M3024-10L
Foetal bovine serum	25 - 50 ml	5 - 10%	Sigma Aldrich	12003C-500ML
L-Glutamine	0.292 g/l	2 mM	Sigma Aldrich	G8540-25G
Sodium Bicarbonate	2.2 g/l	25 mM	Sigma Aldrich	S8540-500G
Penicillin (10 000 units/ml)/ streptomycin (10 000 µg/ml)	5 ml	1%	Sigma Aldrich	P4333-100ML
Mouse embryonic feeder plates				
Mouse embryonic fibroblast (MEF) medium				
Advanced DMEM/F12	450 ml	-	Invitrogen	12634028
Foetal bovine serum	50 ml	10%	Biosera	S04253S181S
L-Glutamine	5.1 ml	2 mM	Invitrogen	25030024
β-Mercaptoethanol	3.5 µl	0.1 mM	Sigma Aldrich	M6250-100ML
Penicillin/streptomycin	5.1 ml	1%	Invitrogen	15140122
Knockout serum replacement (KSR) medium				
Advanced DMEM/F12	400 ml	-	Invitrogen	12634028
Knockout serum replacer	100 ml	-	Invitrogen	10828028
L-Glutamine	5.1 ml	2 mM	Invitrogen	25030024
β-Mercaptoethanol	3.5 µl	0.1 mM	Sigma Aldrich	M6250-100ML
Penicillin/streptomycin	5.1 ml	1%	Invitrogen	15140122
Human induced pluripotent stem cell maintenance and differentiation				
Chemically-defined medium with polyvinyl alcohol				
Ham's F12	250 ml	-	Invitrogen	31765068
IMDM	250 ml	-	Invitrogen	21980065
Chemically-defined lipid concentrate	5.1 ml	1x	Invitrogen	11905031
Monothioglycerol	20 µl	450 µM	Sigma Aldrich	M6145-100ML
Transferrin	250 µl	15 µg/ml	Roche	652202
Insulin	350 µl	7 µg/ml	Roche	1376497
Polyvinyl alcohol	0.5 g	1 mg/ml	Sigma Aldrich	P8136
Penicillin/streptomycin	5.1 ml	1%	Invitrogen	15140122
Human induced pluripotent stem cell differentiation				
Roswell Park Memorial Institute (RPMI) medium				
RPMI Medium 1640 - GlutaMAX™	480 ml	-	Gibco	61870-010
B27 supplement	10 ml	2%	Gibco	17504-044
Minimum essential medium non-essential amino acids (MEM-NEAA)	5 ml	1%	Gibco	11140-035
Penicillin/streptomycin	5 ml	1%	Invitrogen	15140122

Component	Volume	Final concentration	Company	Catalogue number
Hepatozyme-serum free medium (SFM)				
Hepatozyme-SFM	500 ml	-	Invitrogen	17705-021
L-Glutamine	5.3 ml	2 mM	Invitrogen	25030024
MEM-NEAA	10.6 ml	1%	Gibco	11140-035
Chemically-defined lipid concentrate	5.3 ml	1x	Invitrogen	11905031
Transferrin	530 µl	15 µg/ml	Roche	652202
Insulin	742 µl	7 µg/ml	Roche	1376497
Penicillin/streptomycin	5 ml	1%	Invitrogen	15140122
Primary human hepatocytes				
Thawing medium				
Cryopreserved hepatocyte recovery medium	Proprietary formulation		Gibco	CM7000
Incubation medium				
Williams E Medium	500 ml	-	Gibco	A12176-01
FBS	25 ml	5%	Gibco	Primary hepatocyte thawing and plating supplements CM3000
Dexamethasone	50 µl of 10 mM	1 µM		
Human recombinant insulin	500 µl of 4 mg/ml	4 µg/ml		
GlutaMAX™	5 ml of 200 mM	2 mM		
HEPES (pH 7.4)	7.5 ml of 1M	15 mM		
Penicillin/streptomycin	5 ml	1%		
Dissociation buffers and other constituents				
Cell dissociation buffers				
TrypLE express enzyme	-	-	Gibco	12604-021
Cell dissociation buffer	-	-	Gibco	13150-016
Trypsin-EDTA solution	-	-	Sigma Aldrich	T4049-500ML
Gelatine (heated to 56°C for 30 minutes)				
Gelatine	0.5 g	1 mg/ml	Sigma Aldrich	G1890-100G
Water	500 ml	-	Sigma Aldrich	W1503-500ML
Collagenase				
Collagenase IV	500 mg	1 mg/ml	Invitrogen	12634028
Advanced DMEM/F12	400 ml	-	Invitrogen	12634028
Knockout serum replacer	100 ml	-	Invitrogen	10828028
L-Glutamine	5.1 ml	2 mM	Invitrogen	25030024
β-Mercaptoethanol	3.5 µl	0.1 mM	Sigma Aldrich	M6250-100ML
Dispase				
Dispase	0.5 g	1 mg/ml	Gibco	17105-041
Advanced DMEM/F12	400 ml	-	Invitrogen	12634028

Table 2.3: Cytokines used for cell culture

Component	Volume	Final concentration	Company	Catalogue number
Cytokines in feeder co-culture				
FGF-2	4 ng/µl in PBS with 0.1% BSA (w/v)	4 ng/ml	R&D Systems	233-FB-025
Cytokines used during passaging (for thawing and passaging from feeders)				
Y27632 (ROCK Inhibitor)	10 mM in water	10 µM	Sigma Aldrich	Y0503
iPSC maintenance in chemically defined conditions				
Activin A	10 ng/µl in PBS with 0.1% BSA (w/v)	10 ng/ml	R&D Systems	338-AC-005
FGF-2	4 ng/µl in PBS with 0.1% BSA (w/v)	12 ng/ml	R&D Systems	233-FB-025
Endoderm induction				
Activin A	10 ng/µl in PBS with 0.1% BSA (w/v)	100 ng/ml	R&D Systems	338-AC-005
FGF-2	4 ng/µl in PBS with 0.1% BSA (w/v)	80 ng/ml	R&D Systems	233-FB-025
Ly294002 (PI3 kinase inhibitor)	10 µM in DMSO	10 µM	Promega	V1201
CHIR99021 (GSK3 inhibitor)	10 µM in DMSO	3 µM	Stemgent	04-0004
Hepatic specification				
Activin A	10 ng/µl in PBS with 0.1% BSA (w/v)	50 ng/ml	R&D Systems	338-AC-005
Hepatic maturation				
HGF	10 ng/µl in PBS with 0.01% BSA (w/v)	50 ng/ml	Peprotech	100-39
Oncostatin M	20 ng/µl in PBS with 0.01% BSA (w/v)	10 ng/ml	R&D Systems	295-OM-050

Chapter 3: Cell Culturing Techniques

3.1 Materials and Methods

3.1.1 Monolayer Cultures

For 2D monolayer cultures, cells were seeded at densities appropriate for each cell type. Seeding densities did not exceed 2 x 10^4, 5 x 10^5, 2 x 10^6 or 5 x 10^6 cells per 96-well plate, 6-well plate, 25 cm^2 or 75 cm^2 flasks respectively. In the event of inadequate attachment, cells were seeded into Corning CellBIND flasks which have improved hydrophilicity and wettability of the polymer surface.

3.1.2 Three-dimensional Cultures

Three-dimensional cultures vary based on technical complexity and the equipment and consumables required. Adaptability to high-throughput screening (HTS) and biological applicability were key determinants of the method selected for this study. Another consideration was that using scaffold material introduces new potential variables that could be difficult to control when proceeding to the toxicity assays with drugs that could have an affinity to the added matrix. For this reason the hanging drop technique was selected. Perfecta3D® 96-well hanging drop plates were selected and seeding density, size and morphology, medium volume and exchange, aggregation time, humidity control, extracellular matrix (ECM) effects, cell cycle analysis and live-dead staining were assessed.

Agarose (1% (w/v)) was heated at 15 - 30 second intervals in a microwave until completely dissolved and pipetted (1 - 2 ml) into the tray and plate reservoirs to aid in humidity control. HepG2 cells were seeded between 500 - 10 000 cells per well in 50 µl EMEM with 10% FBS. Plates were incubated at 37°C with 5% CO_2 and cells left to aggregate within the drops under gravity. Protein content and approximate dimensions were assessed for up to 21 days. Complete growth medium was replaced every alternate day by removing 10 - 20 µl and replacing 12 - 22 µl. Spheroids were harvested, as required, via air displacement.

3.1.3 Protein Quantitation of Spheroids

HepG2 cell spheroid protein content was quantified using the colorimetric bicinchoninic acid (BCA) assay which is based on the reduction of Cu^{2+} from cupric sulphate to Cu^{+} being proportional to protein content (Smith P, Krohn RI *et al.*, 1985). Spheroids were pooled, centrifuged (1 000 *g*; 5 minutes), washed with PBS and lysed (100 µl per 10 - 20 spheroids) using radioimmunoprecipitation buffer (RIPA: 10 mM Tris-HCL (pH 8), 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5 mM ethyleneglycol tetraaceticacid, 1 % (v/v) Triton X-100, 0.1% (w/v) SDS, 0.1% (w/v) sodium deoxycholate, 140 mM sodium chloride and cOmplete™ protease inhibitor cocktail (Roche Pharmaceuticals)).

Samples were placed in an ultrasonic bath for 10 minutes with 30 second pulses, centrifuged (16 000 *g*; 10 minutes) and supernatants collected for protein quantitation. BCA solution was made by combining a 1:50 ratio of reagent B (4% (m/v) copper II sulfate pentahydrate) to reagent A (2% (m/v) sodium carbonate, 0.16% (m/v) sodium tartrate, 0.9% (m/v) sodium bicarbonate and 1% (m/v) BCA; pH 11.25). A calibration curve was generated (0 - 2 mg/ml) using human serum albumin (HSA). One-hundred and ninety five microliters of BCA solution was added to 5 µl of HSA standards and samples, incubated at 60°C in the dark for 30 minutes and quantified using an ELx800 Bio-Tek microplate reader at a wavelength of 560 nm.

3.1.4 Cell Cycle Analysis

Analysis of cellular DNA content enables subdivision of cells into pre-DNA synthetic interphase (G1), DNA synthesis phase (S), postsynaptic interphase (G2) and mitosis (M). Cellular kinetics of spheroids was done to assess the distribution throughout the cell cycle compared to HepG2 monolayers. Spheroids were cultured for 10 days, collected (n = 20) and washed with pre-warmed PBS. Spheroids were centrifuged (200 *g*; 5 minutes), washed and disaggregated in 1 ml Accutase cell detachment solution at 37°C in a water bath for 45 minutes. Monolayer controls, seeded in 6-well plates, were collected via trypsinization and washed with PBS.

Cells were centrifuged (200 *g*; 5 minutes) and resuspended in 200 µl PBS with 1% FBS. Cells were fixed by adding 2 ml of ice cold 70% ethanol in a drop wise manner under constant agitation. Fixed samples were stored overnight at 4°C. Samples were centrifuged (200 *g*; 5 minutes) and pellets resuspended in 1 ml propidium iodide (PI) staining solution (40 µg/ml PI, 0.1% (v/v) Triton X-10 and 100 µg/ml DNase free RNase) and incubated at 37°C for 40 minutes. Samples were analysed using a Beckman Coulter FC500 flow cytometer at an excitation of 488 nm and emission wavelength of 620 nm.

3.1.5 Fluorescence Microscopy

Live-dead staining was done on spheroids to ensure viability and limited presence of compromised cells. Fluorescein diacetate (FDA) measures esterase activity which serves as a viability probe while PI is a cell-impermeable nucleic acid stain which served as a means of ascertaining membrane integrity. Spheroids were collected at Day 7 and Day 10 (n = 20), centrifuged (200 *g*; 5 minutes) and washed. One millilitre of staining solution (10 µl of 2 mg/ml PI and 2 µl of 5 mg/ml FDA) was added and spheroids were stained for 5 minutes at room temperature, under strict exclusion of light. Spheroids were washed to remove excess stain and visualized using a Zeiss AX10 fluorescence microscope using an excitation wavelength of 488 nm and emission wavelengths of 535 nm (FDA) and 620 nm (PI). Images were processed and analysed using ImageJ software.

3.1.6 Statistics

Protein quantitation data was obtained as absorbance units (0 - 5.0) and a calibration curve used to determine the protein content per spheroid in micrograms (µg). Internal duplicates with a minimum of three independent experiments per culture condition were conducted (n = 6). Cell cycle histograms (n = 3) were analysed using deconvolution software (Wincycle Software; Washington, USA) and subdivided into G1, S, G2 and sub-G1 (background aggregates and debris). Outliers were assessed using the Grubb's test and results expressed as the mean (± standard error of the mean) using GraphPad Prism 5.0. Non-parametric data was analysed using Kruskal-Wallis and p-values of less than 0.05 were considered indicative of significance.

3.2 Results and Discussion

While being easy to use and standardise, monolayer cultures (Figure 3.1) oversimplify the cellular *"bio-mimetic microenvironment"* often diminishing gene expression patterns and inducing phenotypic modulation. Cellular reliance on the microenvironment including both tissue structure and geometry are reported to dictate functionality (Bhadriraju K and Chen CS, 2002; Peters TS, 2005).

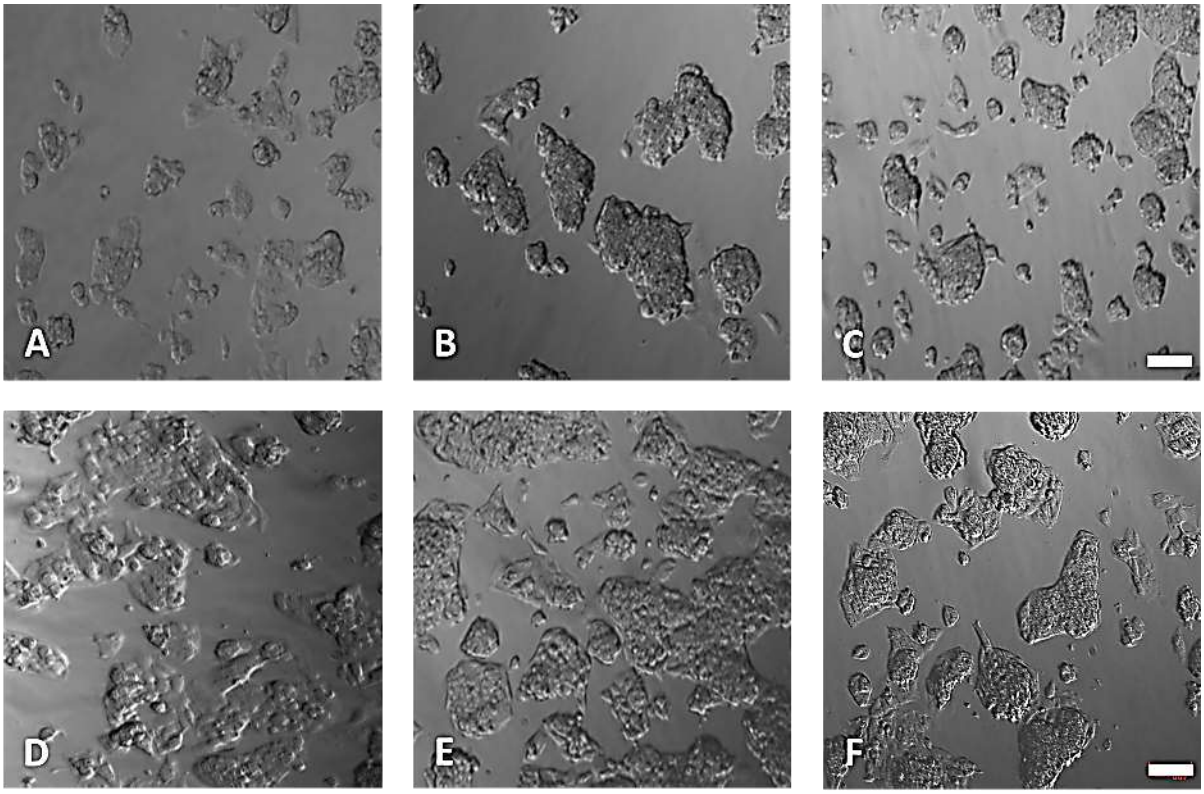

Figure 3.1: Phase contrast images (Zeiss inverted Axiovert CFL40 microscope) of A - C) HepG2 monolayers cultures in EMEM supplemented with 5% FBS and D - F) HepG2 monolayers cultures in EMEM supplemented with 10% FBS. Scale bar: 100 µm.

Mimicking higher order processes requires assembly of functional hierarchical structures (100 µm - 1 mm) using single cell (~10 µm) building blocks. Use of hepatocytes or hepatocyte-like cells in 3D culture enhances the ability to generate accurate miniature organoids (Khetani SR and Bhatia SN, 2008). Self-aggregated rat primary hepatocytes cultured in dynamic rotating-wall vessels (RWV) express enhanced hepatocyte-specific synthetic and metabolic functions. Global gene expression analysis has established a major role for HNF4α in maintaining hepatic functionality in spheroids (Chang TT and Hughes-Fulford M, 2014). However, the choice of 3D culture impacts the degree of hepatic functionality with dynamic RWV cultures recapitulating bile duct-like and blood vessel-like structures more efficiently than hanging drops or spinner flasks (Cai J, DeLaForest A *et al.*, 2012).

Support for genotypic and phenotypic maintenance and maturation of hepatocytes in 3D culture is vast. However, some techniques are limited in scalability and high-throughput. A static 3D culture format, amenable to high-throughput, capable of producing homogenous, size controlled spheroids was chosen. Hanging drops (Figure 3.2) generate scaffold and extracellular matrix free spheroids which enhance cell-cell communication while being easy to extract and process for downstream applications. During optimisation, droplet volume was varied with stable droplets of 50 µl, the upper limit recommended by manufacturers, originally chosen as the seeding volume. Due to various factors such as spreading medium (Figure 3.3) and subsequent loss of spheroids, seeding volume was decreased to 45 µl for further experimentation.

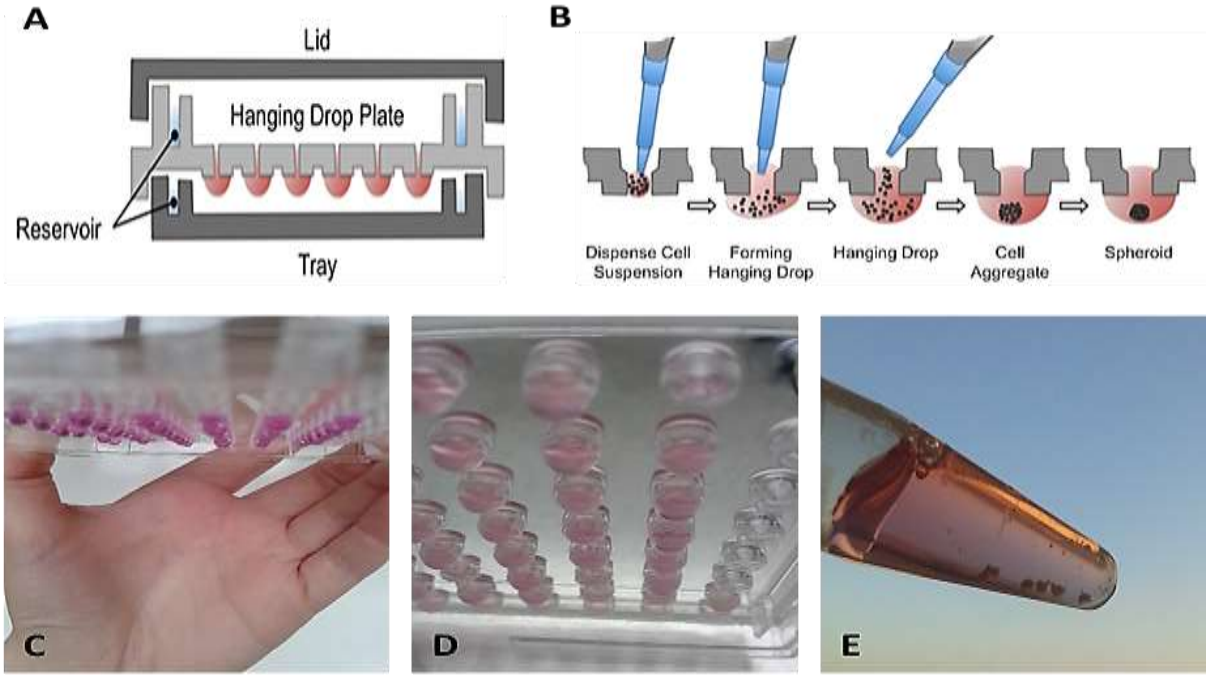

Figure 3.2: A - B) Schematic of the general plate format and process for cell seeding and spheroid formation, C - D) Photographs of HepG2 cells seeded in hanging-drop plates and E) Photographs of HepG2 spheroids collected in an Eppendorf tube at Day 15.

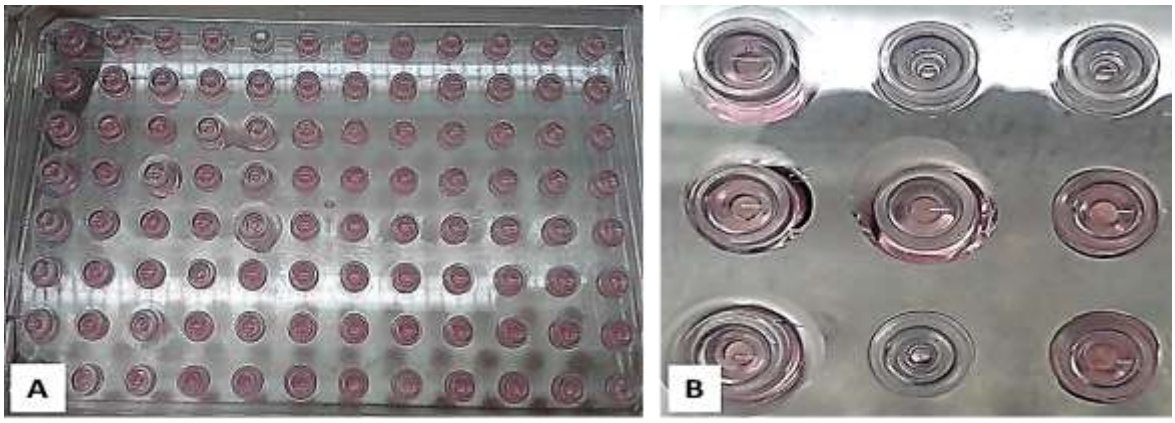

Figure 3.3: Photographs of A) Cells seeded in a hanging drop plate and B) A 96-well plate illustrating the medium spreading around the access hole which resulted in merging with adjacent droplets.

HepG2 spheroids, seeded at 5 000 cells, produced approximately 6.62 (± 0.56) micrograms of protein per spheroid. This increased 3-fold by Day 10 to 21.18 µg and plateaued at approximately 26 µg by Day 21. The most notable increase (~10 µg) was seen between Day 3 and Day 8. Improved seeding, protein pooling and sample preparation yielded more consistent and reproducible results with protein content of 3.75 (± 0.29) µg/spheroid at seeding, increasing to 15 µg/spheroid at Day 14.

Seeding density was varied to ascertain various morphological features. While multiple layers of cells were evident at all seeding densities, spheroids were not concentric in multiple dimensions. Characterisation of spheroid morphology of breast and prostate cell lines produce four distinct morphologies namely; round, mass, grape-like and stellate (Härmä V, Virtanen J *et al.*, 2010; Kenny PA, Lee GY *et al.*, 2007). While mostly appearing round, morphological characteristics cannot be directly compared to hanging drops where gravitational forces ensure a measure of cell aggregation. Progress in tissue engineering has improved the diversity of 3D culture models. The term spheroid is used inconsistently throughout literature and is sometimes used to refer to loosely packed aggregates which can't be manipulated or transferred. These cultures often lack cell-cell and cell-matrix interactions as well as spherical geometry which alters the pathophysiological gradients of classically characterised spheroids (Hirschhaeuser F, Menne H *et al.*, 2010).

Extracellular matrices such as collagen, laminin or methylcellulose are known to effect morphology. Clustering and compaction of hepatoma cell lines (HepG2/C3A, PLC/PRF/5, Hep3B) in hanging drops illustrate three assembly stages. Initial aggregation with strong reliance on integrins, a delayed period for cadherin accumulation and a final compaction stage (Lin R-Z, Chou L-F *et al.*, 2006). At 24 hours, cell-specific differences in compactness are discernible (Figure 3.4) resulting in categorization as compact spheroids, tight aggregates and loose associations of cells (Ivascu A and Kubbies M, 2007; Lin R-Z, Chou L-F *et al.*, 2006).

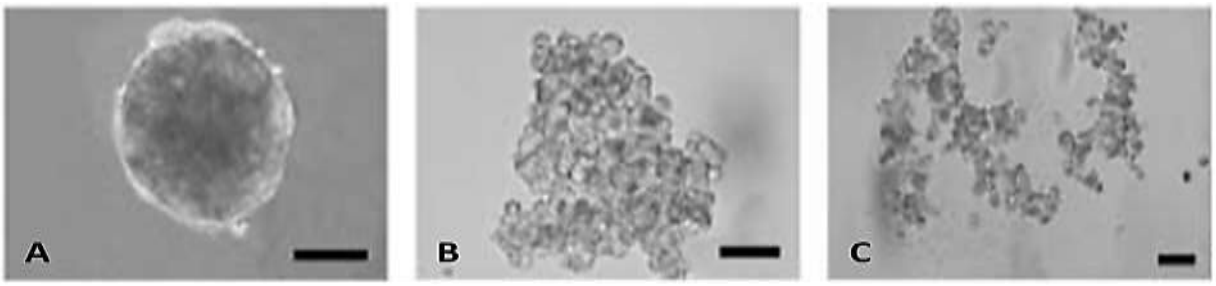

Figure 3.4: Comparative analysis of spheroid-forming capability among three hepatoma cell lines. A) Hep3B spheroids with a smooth surface, B) HepG2C3A cells forming loose cellular aggregates and C) PLC/PRF/5 cells with no obvious aggregation. Scale bar 200 µm. (Reprinted with permission from Springer, license number 3961450962686, (Lin R-Z, Chou L-F *et al.*, 2006))

Morphologically, HepG2 spheroids appeared to be tight aggregates as opposed to compact spheroids like Hep3B cell spheroids which exhibit high levels of both E-cadherin and N-cadherin. Collagen promotes integrin activity and markedly increases compaction during initial culture phases (0 - 18 hours). Differential expression of E-cadherin and integrins, influenced by cell type and ECM, alter spheroid shape (Lin R-Z, Chou L-F *et al.*, 2006).

In support of classification as compact spheroids, tight aggregates and loose associations of cells, it is suggested that a single cadherin (E- or N-cadherin) or ECM-integrin receptor is responsible for tight cell packaging. E- and N-cadherin expressing cells can spontaneously aggregate whereas surface integrin ß1-ECM plays a role in compaction. Furthermore, transition from aggregates to compact spheroids can be achieved with non-toxic medium supplement such as reconstituted basement membrane (rBM: Matrigel)(Ivascu A and Kubbies M, 2006).

Attempting to manipulate compactness and shape was done using collagen and methylcellulose as additives to culture medium. Acid-soluble collagen was used, which while having minimal effect on pH, decreased protein content. Protein content, used as a surrogate for inferring cell enumeration, was 7.89 (± 0.21) µg/spheroid as opposed to 21.18 (± 0.29) µg/spheroid where no collagen was added.

Methylcellulose, a chemical derivative of cellulose, is suggested to enhance compaction of spheroids and increase medium viscosity thereby stabilizing droplets. Cells were seeded in EMEM containing 0.5%, 1% and 2% methylcellulose and monitored for morphology and protein content. Protein content of spheroids cultured in methylcellulose-based spheroids paralleled that in EMEM alone over 10 days but illustrated varied morphologies (Figure 3.5). Increasing viscosity negated gravitational forces which would normally allow cells to settle to the droplet apex.

Consequently, at higher concentrations of methylcellulose, multiple smaller spheroids formed as opposed to a single, concentric spheroid, increasing the heterogeneity of the technique. While adding to droplet robustness, use of methylcellulose at these concentrations did not appear to promote compaction. Due to the limited advantages conferred by the tested extracellular matrices, neither methylcellulose nor collagen were included during initial spheroid formation (Day 0 – Day 4) in hanging drops.

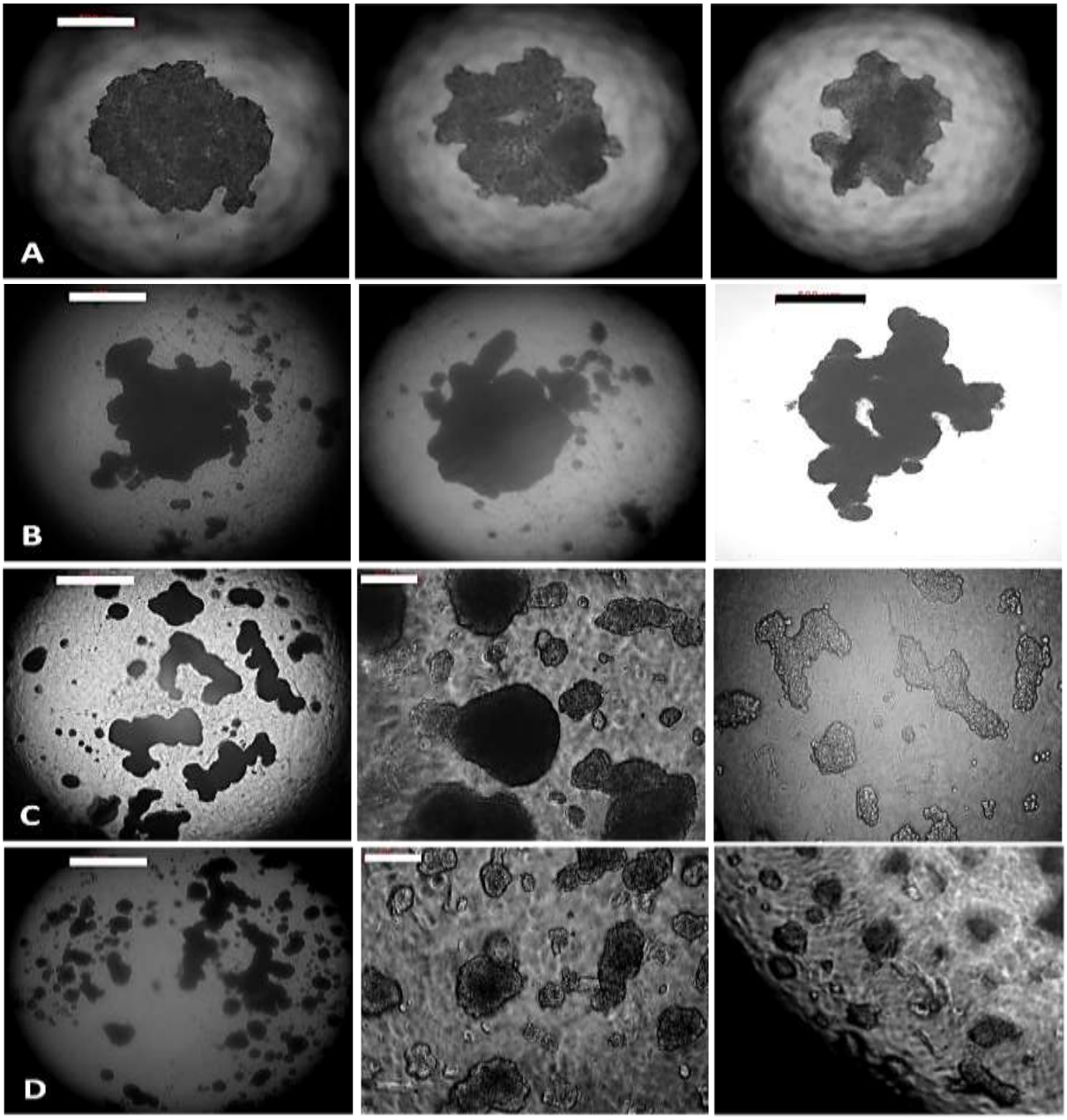

Figure 3.5: Phase contrast images (Zeiss inverted Axiovert CFL40 microscope) of HepG2 cell cultured in hanging drops with variable medium additives. A) EMEM, B) EMEM with 0.5% methylcellulose, C) EMEM with 1% methylcellulose and D) EMEM with 2% methylcellulose. Scale bar: 500 µm or 100 µm (images with no scale bar are scaled as per the image to the left).

Cavanar *et al.* demonstrated that some spheroids in hanging drops have morphology that curves with the droplet radius and do not form concentric spheroids (Figure 3.6). This shape may confer an advantage as transport gradients are not as limited which reduces the risk of central necrosis in spheroids of 500 - 600 µm (Cavnar SP, Salomonsson E *et al.*, 2013). Based on these parameters and the inability of other matrices to confer significant advantages, the natural morphology of HepG2 cells as tight aggregates was considered sufficient to be termed spheroids to be used for experimentation.

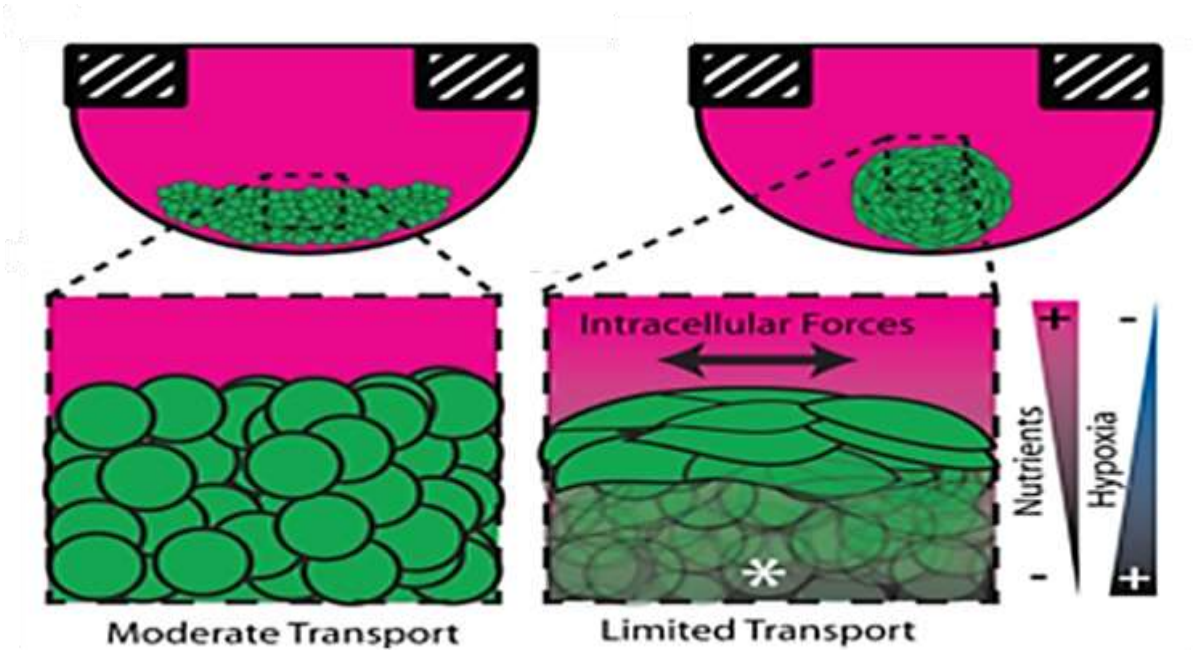

Figure 3.6: Schematics of spheroid morphology. Left: spheroids which curves with the droplet radius and have loose interactions between cells, enhancing permissiveness to nutrient transport. Right: spheroids which appear more spherical with tight intracellular forces limiting transport gradients. (Reprinted from open access journal with permission, (Cavnar SP, Salomonsson E *et al.*, 2013))

Spheroids, whether tight aggregates or completely compact, can be divided into distinctive zones. Zonation includes; an outer proliferation zone, intermediate quiescent zone of viable cells and a potential inner core where cells begin undergoing necrosis. Based on spheroid size and shape, characteristics which can be altered include nutrient and waste gradients, ease of compound penetration, degree of hypoxia and general ability to mimic *in vivo* tissue types (Figure 3.7) (Mehta G, Hsiao AY *et al.*, 2012).

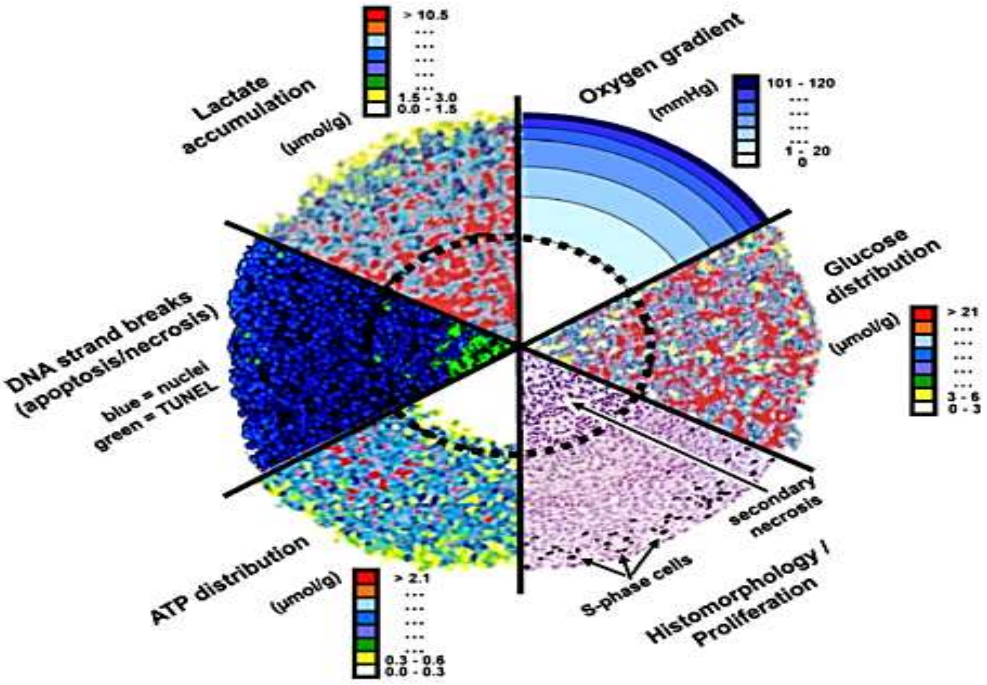

Figure 3.7: Combination of analytical images of spheroid median sections studied with different technologies. Together these enable the concentric arrangement of cell proliferation, viability and the micromilieu in large spheroids to be understood. (Reprinted with permission from Elsevier, license number 3961461153291, (Hirschhaeuser F, Menne H *et al.*, 2010))

Comparing the protein content of seeding densities of with 5 000 and 10 000 cells was done up to 14 days. An assumption was that approximate protein content would double for the higher seeding density, but this was not evident. The distinctive zonation of spheroids described could be responsible for this occurrence. Limitations in the outer proliferation (Figure 3.8) zone size may suggest that at higher seeding densities, a larger quiescent zone develops more rapidly thereby slowing the cells doubling time resulting in marginal protein difference from the different seeding densities.

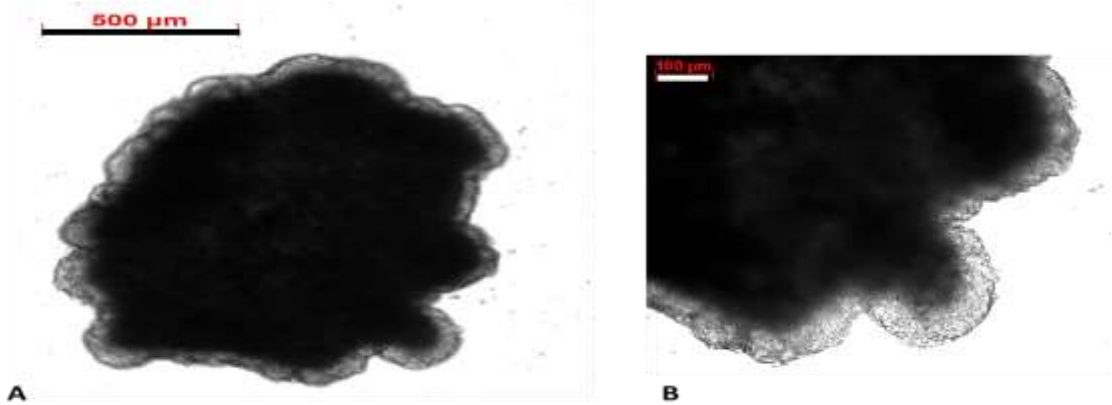

Figure 3.8: Phase contrast images (Zeiss inverted Axiovert CFL40 microscope) of A) HepG2 spheroids (Scale bar: 500 µm) and B) Higher magnification of the same spheroid with the more transparent periphery suggestive of the proliferation zone. Scale bar: 100 µm.

To promote amenability to high-throughput, ease of culture approaches were considered. Due to laborious medium exchange, spheroids were collected in a 6-well plate with 1% methylcellulose in the medium. Enhanced viscosity prevented agitation, breaking or adhering to culture surfaces. Spheroids dropped from Day 4 to 7 were well maintained as individual, homogenous entities (Figure 3.9).

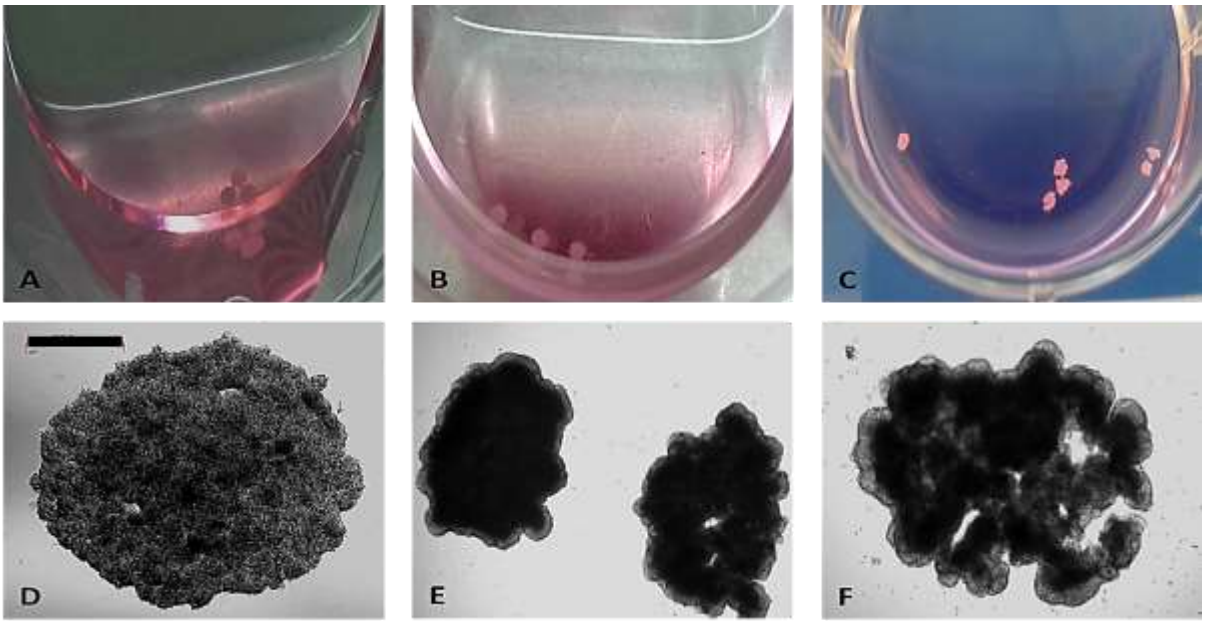

Figure 3.9: Images of HepG2 spheroids in methylcellulose-based medium. A - C) Photographs of spheroids dropped into a 6-well plate from Day 7, D - E) Phase contrast image of HepG2 spheroids in methylcellulose-based medium at Day 7 to Day 10 and F) HepG2 spheroids in EMEM at Day 10 which have been damaged due to plate agitation. Scale bar of 500 µm.

Once optimized and heterogeneous spheroids were generated and easy to manipulate, additional characterisation was done. Protein content was used as an inference for cellular proliferation but assessing spheroid viability throughout the time-course of culture was essential. This was determined through fluorescence microscopy to ascertain viability and membrane integrity. Additionally, cell cycle analysis was done to ensure that cells were still dividing corresponding to the proliferation zone. And further to establish the degree of quiescence (G1 cell cycle arrest) and cell death (sub-G1 cell cycle).

Cell cycle analysis of disaggregated HepG2 spheroids (Figure 3.10) patterned the cell cycle of unsynchronized monolayers with an observable increase in the G1 phase. This could be due to the notion of spheroid zonation with a small margin of proliferating cells being present in the outer zone. The minimal sub-G1 phase demonstrates that cells were still viable when cultured for up to 10 days from an initial seeding density of 10 000 cells per spheroid. Fluorescence microscopy (Figure 3.11) results corroborated cell viability with homogenous FDA staining throughout spheroids and minimal non-specific staining of PI demonstrating no presence of a central necrotic core.

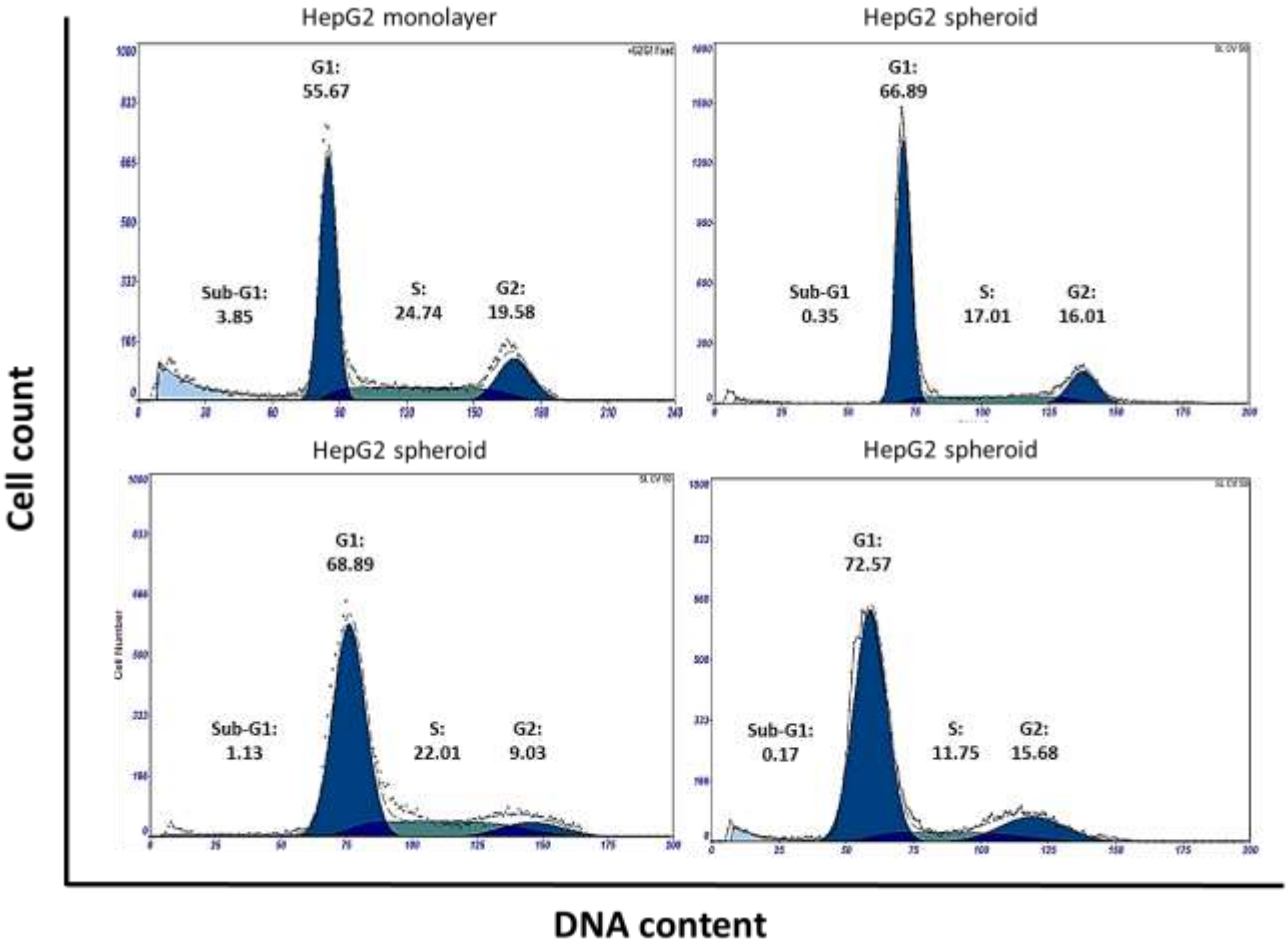

Figure 3.10: Cell cycle analysis of HepG2 monolayers and spheroids. Deconvolution divided histograms into sub-G1 (cell debris and DNA fragments), G1 phase, S phase and G2/M phases with the associated percentage across replicates (n = 3).

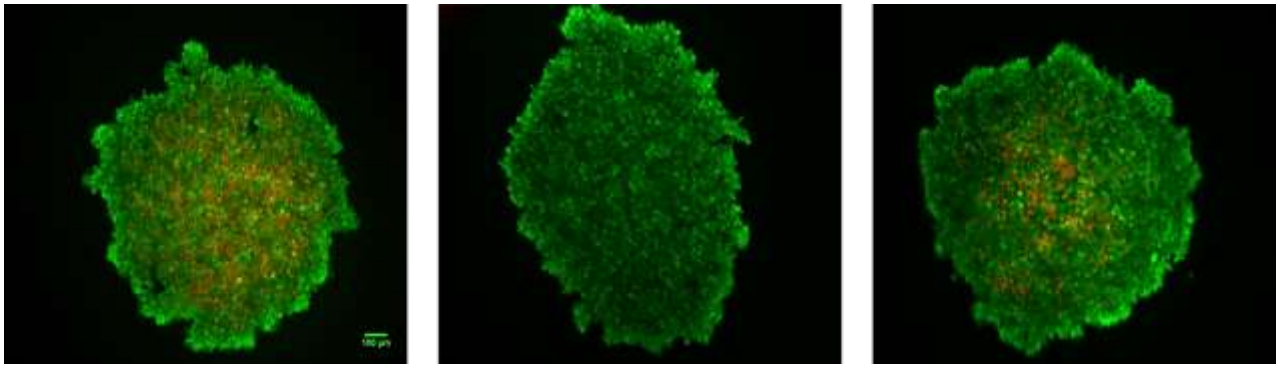

Figure 3.11: Fluorescence microscopy images of HepG2 cell spheroids stained with FDA illustrating cell viability and non-specific PI staining indicative of cell membrane permeability. Scale bar: 100 µm.

Prior to any proteomic or cytotoxicity experiments, a new HepG2 cell line cell line with known passage number was obtained. Spheroids for experimentation were generated at a seeding density of 10 000 cells decreased in 45 µl of medium. This combination of conditions, resulted in fewer spheroids being lost, and generated more reproducible and compact spheroids (Figure 3.12). As tumour cells aberrantly express adhesion molecules, it is possible that aggregation and compaction may differ with passage number as was observed using different batches of HepG2 cells. If required, reduced or inconsistent compaction may be salvage with use of rBM (Ivascu A and Kubbies M, 2006). This was however not required downstream with passage number maintained below 30.

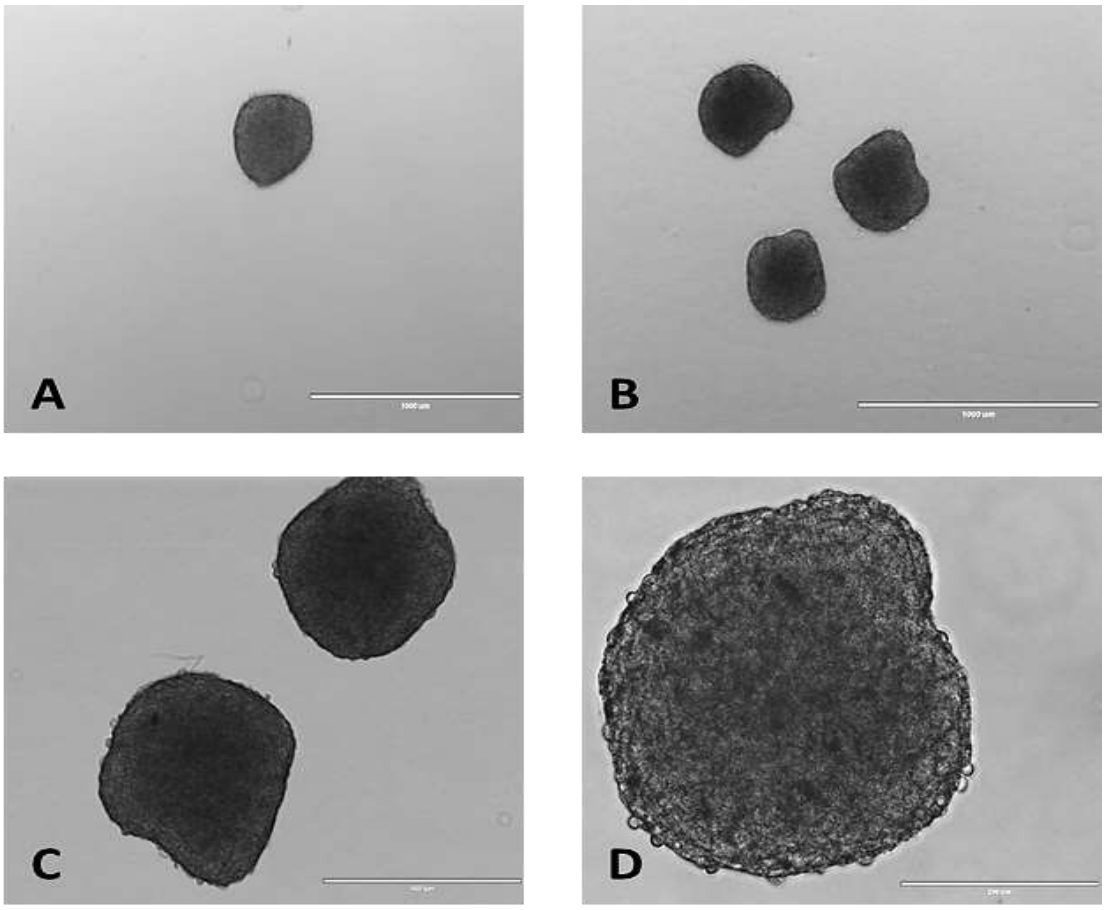

Figure 3.12: Phase contrast images (EVOS FL Cell Imaging System) of spheroids at Day 10 taken in a 6-well plate at various magnifications A - B) Scale bar: 1000 µm, C) Scale bar: 400 µm and D) Scale bar: 200 µm.

3.3 Conclusion

Important features of producing spheroids include applicability to the biological question, spheroid homogeneity, sufficient protein for downstream applications, the ability to transfer and amenability to high-throughput. Hepatocytes in 3D are capable of mimicking higher order processes with functional hierarchical structures essential in systems communications. The advantages of using hanging drops observed in these experiments were the ease of manipulation, homogeneity, ability to collect multiple spheroids as well as there being no requirement for removal of or extraction from complex extracellular matrices.

Optimisation resulted in spheroids of 10000 cells in 45 µl EMEM with 10% FBS cultured up to 10 days being chosen for downstream applications. Data presented in this chapter illustrates that while this method is easy to adopt in its simplest format, it can be manipulated to produce variants of spheroids which are both more or less desirable. In theory, 3D cultures should express more markers associated with mature phenotypes. Whether maturation is based on altering the hepatic phenotype or rearrangement of morphological features was yet to be established. Detailed analysis of the proteomic perturbations induced by spheroidal HepG2 cultures are discussed in Chapter 7.

Recent advances demonstrating equivalence to hanging drops include Corning® spheroid microplates which have altered well geometry and ultralow hydrophobic, non-ionic attachment surfaces. This enables cells to cluster unattached in the bottom of the well, which could potentially be used in HTS to reduce the laborious features associated with hanging drops.

Chapter 4: Characterisation of Proteome Changes during Hepatocyte Differentiation

4.1 Materials and Methods

4.1.1 Protein Collection

Cells harvested on different days throughout differentiation, from biological replicates, were washed with PBS to remove any soluble media derived protein and lysed using protease-inhibitor containing RIPA buffer. Cells lysis was assisted by ultrasonic disruption in an ultrasonic bath for 10 minutes with 30 second pulses, centrifuged (16 000 *g*; 10 minutes) at 4°C and the supernatant collected. Supernatants were quantified for protein concentration using a BCA assay (Chapter 3: Protein Quantitation of Spheroids) and assayed using gel-based and mass spectrometry-based proteomics.

4.1.2 Gel-based Proteomics

Proteins were separated by mass using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Twenty micrograms of protein was mixed in a 1:1 ratio with Laemmli sample buffer (two-times sample buffer: 0.125 M Tris-HCL (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.004% (w/v) bromophenol blue). Samples were heated to 95°C for 5 minutes, cooled, clarified by centrifugation and loaded onto precast Mini-PROTEAN TGX polyacrylamide gels (4 - 15%; 10-wells, 30 µl). Proteins were separated (running buffer: 0.1% (w/v) SDS, 25 mM Tris-base and 19.2 mM glycine) using a Mini-PROTEAN Tetra System at 80 V for 15 minutes and 160 V until completion. Gels were fixed (45% (v/v) methanol, 1% (v/v) acetic acid in ddH_2O) for 30 minutes and stained overnight using Coomassie brilliant blue solution (0.1% (w/v) Coomassie brilliant blue R250, 50% (v/v) methanol, 10% (v/v) acetic acid in ddH_20). Destaining was done with dH_2O under agitation until bands were clearly visible and gels scanned using a Bio-Rad Gel-Doc EZ Imager.

4.1.3 Mass Spectrometry-based Proteomics

4.1.3.1 Protein Preparation and Digestion for Tandem Mass Tagging

Accurately aliquoted amounts of 50 µg of cell derived protein, from individual differentiations (without pooling), was volume adjusted to 100 µl using solubilisation buffer (0.1% (w/v) SDS, 50 mM HEPES; pH 7.4). Samples were reduced using dithiothreitol (DTT; 10 mM in 100 mM HEPES; pH 7.4) for 1 hour at 37°C and alkylated with iodoacetamide (IAA; 25 mM in 100 mM HEPES; pH 7.4) for 2 hours at room temperature under strict exclusion of light. Proteins were precipitated overnight at 4°C with 10-volumes of pre-chilled acetone. Following centrifugation (16 000 *g*; 10 minutes at 4°C), the acetone layer was aspirated and discarded, pellets dried and resuspended in 100 µl of 100 mM HEPES (pH 8.5).

Sequence-grade modified trypsin (1.25 µg trypsin per 50 µg protein) was added and samples incubated in a water bath at 37 °C for 1 hour. A second trypsin aliquot of equivalent mass was added and samples digested overnight in a water bath at 37°C.

4.1.3.2 Peptide Tandem Mass Tag Labelling

Six-plex tandem mass tags were used for stable isotope labelling of the peptides for quantitative comparison of the proteome at different stages of differentiation (Table 4.1). Tags were equilibrated to room temperature and solubilized by adding 41 µl mass spectrometry-grade acetonitrile (25 - 30% organic content per 0.8 mg for 100 µl sample). Peptides from digested samples were clarified by centrifugation (16 000 *g*; 10 minutes at 4°C), added to specified tags for individual labelling and incubated for 2 hours at room temperature under constant agitation. Labelling was quenched for 1 hour at room temperature under constant agitation with 8 µl of 5% hydroxylamine (100 mM HEPES; pH 8.5). Samples were further quenched overnight at 4°C with 100 µl of dH_2O. Differentially mass labelled samples were combined, vacuum centrifuged (10°C) to dryness and stored at -80°C.

Table 4.1: Tandem mass tags used for HLC differentiation time courses

	Differentiation time-course from Day 1 - Day 35 (HLCTC)		Hepatocyte maturation time-course from Day 16 - Day 40 (HLCLTC)	
TMT label	**Replicate 1**	**Replicate 2**	**Replicate 1**	**Replicate 2**
126	iPSC (Day 1)	iPSC (Day 1)	HLC (Day 36)	Late-HLC (Day 40)
127	Endoderm (Day 5)	Endoderm (Day 3)	Maturation (Day 32)	HLC (Day 36)
128	Progenitors (Day 10)	Endoderm (Day 5)	Maturation (Day 28)	Maturation (Day 32)
129	Maturation (Day 25)	Progenitors (Day 7)	Maturation (Day 24)	Maturation (Day 28)
130	Maturation (Day 30)	Progenitors (Day 10)	Maturation (Day 20)	Maturation (Day 20)
131	HLC (Day 35)	Maturation (Day 30)	Maturation (Day 16)	Maturation (Day 16)

4.1.3.3 Solid Phase Extraction

Solid phase extraction (SPE) was performed as a clean-up step to remove non-volatile buffers, dry salts and unreacted tags. SepPak C18 cartridges (100 mg) were conditioned using 1 ml each of acetonitrile followed by 0.05% acetic acid (pH 3) and dH_2O with 0.1% trifluoroacetic acid (TFA; pH 1). Pooled TMT labelled peptides, resolubilized in 800 µl of 0.1% TFA, were loaded slowly onto conditioned cartridges and washed using 1 ml each of 0.1% TFA and 0.05% acetic acid. Peptides were eluted using 70% acetonitrile with 0.05% acetic acid. A final elution was done using 1 ml pure acetonitrile. Peptide-containing eluents were vacuum dried (10°C) and stored at -80°C prior to UPLC fractionation.

4.1.3.4 Reverse Phase High Performance Liquid Chromatography

4.1.3.4.1 Instrument Parameters and QC

Reverse phase UPLC (RP-UPLC) was done using a Waters ACQUITY UPLC system coupled to a photodiode array (PDA) detector. Priming was done for 2 minutes using 50:50 mobile phases A and B (Table 4.2) at a flow rate of 4 ml/minute. Samples were separated using an ACQUITY UPLC Bridged Ethyl Hybrid (BEH) C18 column (130Å, 1.7 µm, 2.1 x 150 mm, trifunctionally-bonded alkyl). Mobile phases A and B (50:50) were run through (0.125 ml/minute to 0.244 ml/minute) the column and pressure left to stabilize below 10 000 psi (delta value less than 50) before reducing to 5% mobile phase B. Instrument setup and mobile phases are summarised in Table 4.2

Table 4.2: Waters ACQUITY system conditions and settings for RP-UPLC

Binary solvent manager	
Solvent A (aqueous)	20 mM ammonium formate, pH 10
Solvent B (organic)	80% acetonitrile 20 mM ammonium formate, pH 10
Pressure limit: low	0 psi
Pressure limit: high	13 000 psi
Seal wash	5 min (11 min run) 10 min (75 min run)
PDA detector	
Wavelength range	210 - 400 nm
Sampling rate	20 points per second
Filter time constant	Normal (0.1 seconds)
Resolution	1.2 nm
Sample manager	
Autosampler temperature	10°C
Column temperature	40°C
Sample loop option	Partial loop
Wash solvents	Weak - 5:95; AcN:H_2O (500 µl) Strong - 50:50; AcN:H_2O (500 µl)

Three column washes were conducted using 11 minute run times (Table 4.3). Column and system quality control (QC) was done using 1 nmol cytochrome C (Figure 4.1) over the same 11 minute run. After successful QC results were obtained, two additional column washes were done prior to loading and fractionation of the sample.

Table 4.3: Mobile phase gradient for wash steps and cytochrome C standard using RP-UPLC

Time (Minutes)	Flow (ml/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	0.244	95	5
7.00	0.244	40	60
7.10	0.244	5	95
8.50	0.244	5	95
8.60	0.244	95	5

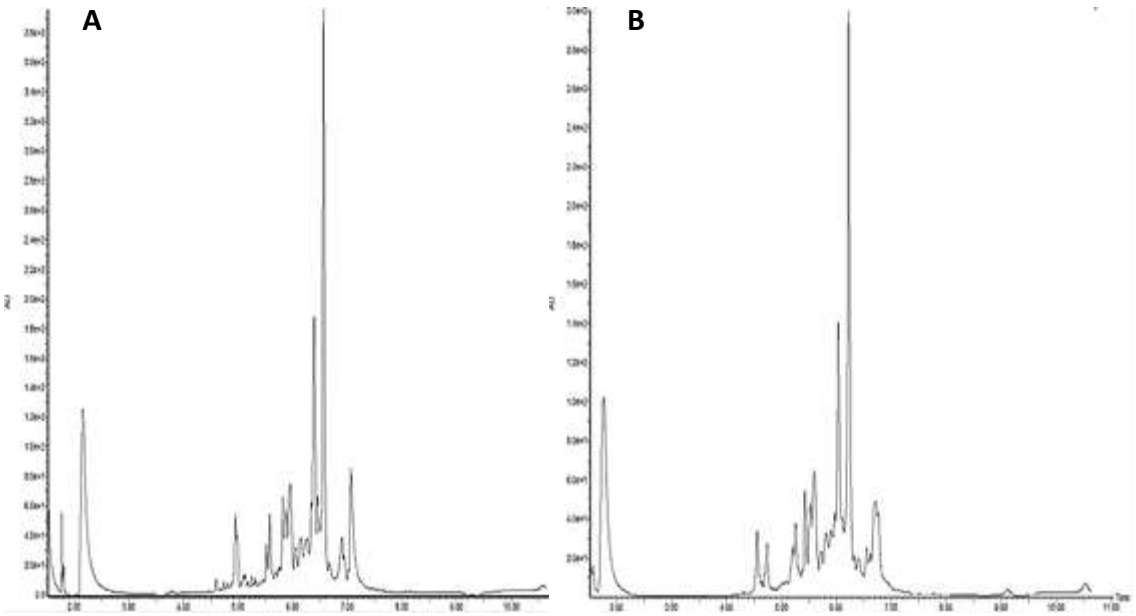

Figure 4.1: A) Reference quality control standard of cytochrome C and B) Cytochrome C standard run prior to sample loading and fractionation.

4.1.3.4.2 Sample Peptide Fractionation

Peptides samples were resuspended in 100 µl of 4% acetonitrile with 20 mM ammonium formate (pH 10). Ninety-five microliters was loaded, via a single partial loop injection, onto a Waters ACQUITY UPLC BEH C18 column. Peptides were eluted at a flow rate of 0.244 ml/minute using an initial isocratic low organic phase followed by a 50 minute linear gradient of increasing percentage mobile phase B and a final 5 minute clean-up step before re-establishing the starting conditions (Table 4.4). The total run time was 75 minutes which included the re-equilibration. Chromatographic separation was monitored between 200 and 400 nm using a PDA detector. Residual salts eluted during the initial 10 minute isocratic stage (Figure 4.2B) after which the majority of peptides eluted during the gradient stage. Peptide containing fractions eluting between 21 and 56 minutes were collected at variable intervals (Table 4.5).

Table 4.4: Mobile phase gradient for peptide elution using RP-UPLC over a 75 minute run

Time (Minutes)	Flow (ml/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	0.244	95	5
10.00	0.244	95	5
60.00	0.244	25	75
62.00	0.244	0	100
67.50	0.244	0	100
67.60	0.244	95	5

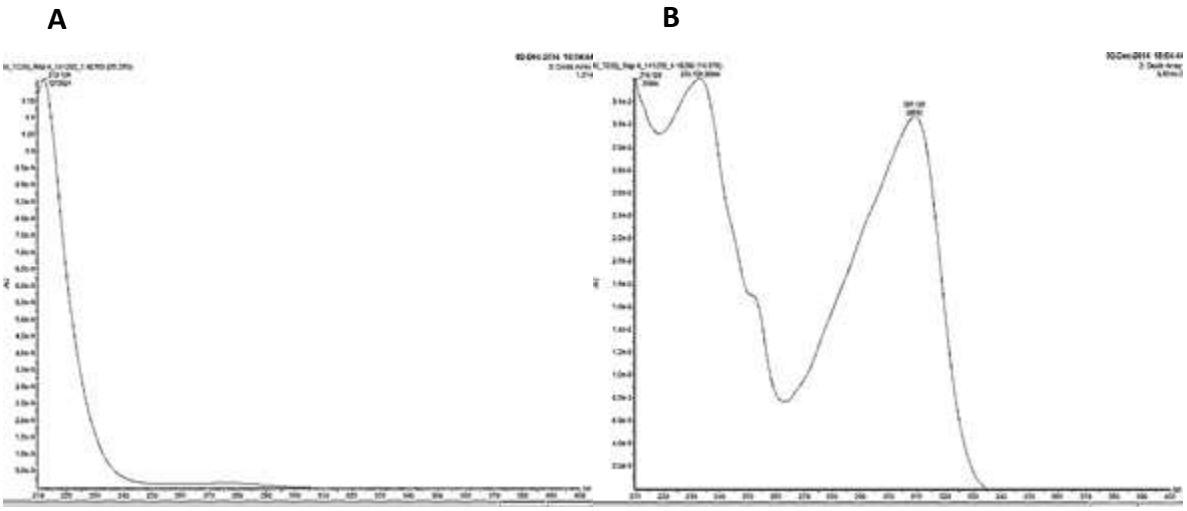

Figure 4.2: Diode array trace for A) Typical protein spectra with a single smooth peak from 210 nm and B) Atypical spectra with multiple peaks and shoulders indicative of salt or other contamination.

Table 4.5: Fraction collection of TMT-labelled peptides

Time span	Duration of collection (Minutes)	Number of fractions
0 - 20	5	4
21 – 26	2	3
27 – 50	1	24
51 - 56	2	3
57 - 66	5	2

Fractions were immediately vacuum dried (10°C), as high pH can deaminate amino acids, and stored at -80°C. To reduce MS run time, selected hydrophobic and hydrophilic fractions were pooled to provide 15 fractions having approximately the same peptide abundance but spread across the gradient from 21 to 56 minutes (Figure 4.3). Pooled samples were resuspended in a total of 30 µl of 0.1% formic acid, as the loading solvent, and transferred to conical-bottom auto-sampler vials for liquid chromatography tandem-mass spectrometry (LC-MS/MS) analysis.

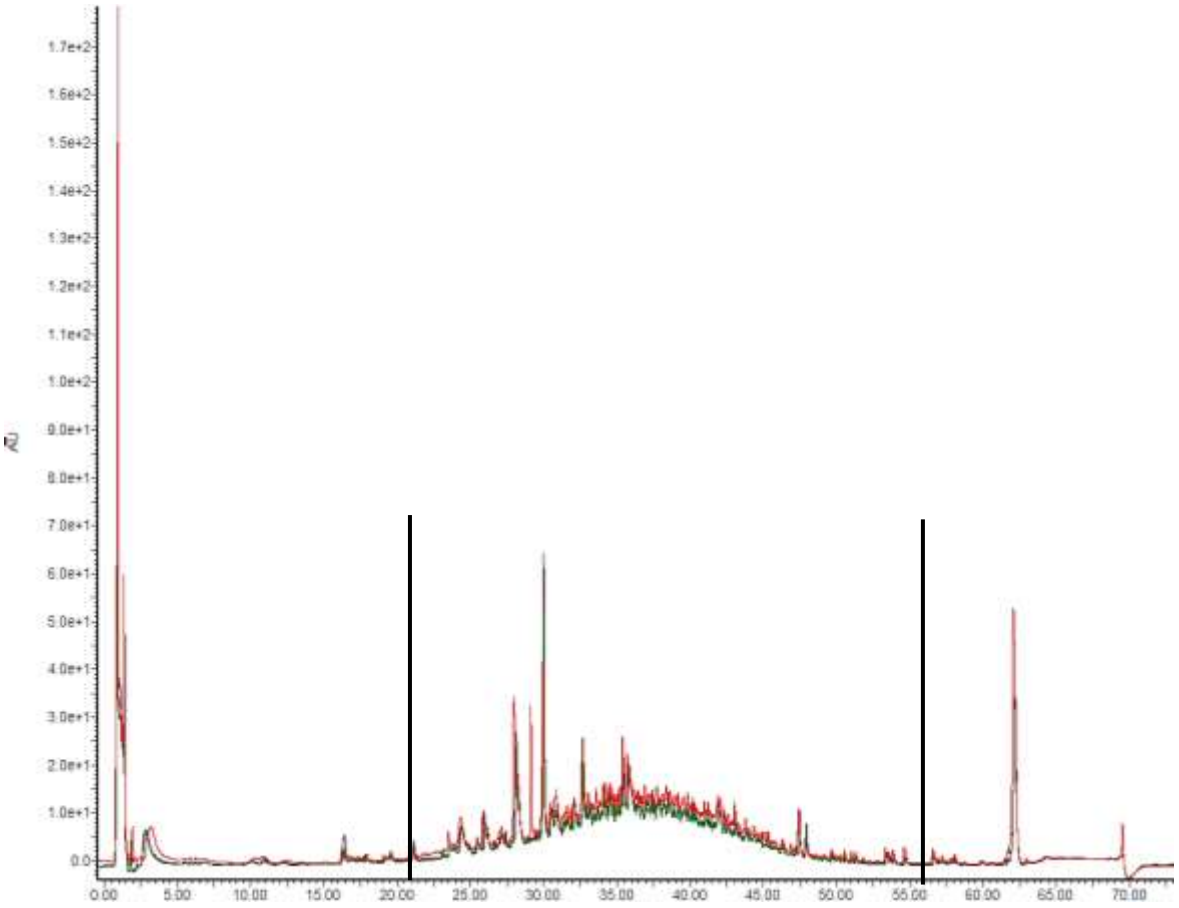

Figure 4.3: Overlay chromatogram of peptide separation from three different samples, where hydrophilic peptides with affinity for the mobile phase eluted prior to the hydrophobic peptides interacting with the hydrophobic stationary phase. Black lines indicated the time frame for collecting peptide containing fractions.

4.1.3.5 LC-MS/MS Analysis

Samples were analysed at the Cambridge Centre for Proteomics Core Facility using a Dionex Ultimate 3000 RSLCnano LC (Thermo Scientific; Massachusetts, USA) system and a Thermo Scientific Q Exactive (QE) Hybrid Quadrupole-Orbitrap Mass Spectrometer. Peptides (1 - 2 µg) were injected onto an Acclaim PepMap 100 C18 pre-column (Thermo Scientific; 100Å, 5 µm, 300 µm x 5 mm) using an Ultimate 3000 auto-sampler with 0.1% formic acid for 3 minutes at a flow rate of 10 µl/minute.

Peptides were eluted onto a PepMap C18, EASY-Spray LC analytical column (Thermo Scientific; 100Å, 2 µm, 75 µm x 500 mm). Peptide separation used a flow rate of 300 nl/minute of mobile phase A (MS-grade H_2O with 0.1% formic acid) and mobile phase B (80% acetonitrile with 0.1% formic acid). A linear gradient from 4 - 40% mobile phase B over 100 minutes was used for separation using a 120 minute run time (Table 4.6).

Table 4.6: Dionex Ultimate 3000 RSLC nanoUPLC system mobile phase gradients and valve position

Time (Minutes)	Flow (µl/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	Valve position: 6_1		
0.00	0.300	96	4
3.00	MS start; Valve position: 1_2		
5.00	0.300	96	4
10.00	Inject valve to load		
10.50	Wash		
15.00	Inject valve to inject		
60.00	Loading pump pressure acquisition off		
100.00	0.300	60	40
100.30	0.300	10	90
110.00	0.300	10	90
110.30	0.300	96	4
119.00	Valve position: 6_1		
120.00	0.300	96	4

LC eluent was introduced into the mass spectrometer via a Thermo Scientific Easy-Spray Source which provides integrated, nano-electrospray ionization. Mass-to-charge (m/z) ratios of positively charged ions was measured in the Orbitrap mass analyser (resolution: 70 000). Data dependent scans of the top 20 most abundant ions, with charge states between 2+ and 5+, were automatically isolated. Selected ions were fragmented by higher energy collisional dissociation (HCD) in the quadrupole mass analyser and fragments measured in the Orbitrap (resolution: 17 500). Further details of experimental and instrumental parameters are outlined in Table 4.7. Data was collected as raw data files that were exported directly in a format that could be used with bioinformatics analysis software.

Table 4.7: QE experimental parameters for full MS / data-dependent (dd) -MS²

Global Settings (Overall method [TopN: two event scans])	
Lock mass (real-time recalibration by correction of m/z shifts arising from instrumental drift)	If all present (2 entries: mass and polarity): 445.12000, positive 519.13860, positive
Chromatogram peak width (median of peak widths for all identified peptides; full width at half maximum [FWHM])	25 s
General	
Runtime	0 - 120 minutes
Polarity	Positive
In-source collision-induced dissociation (CID)	0.0 eV
Default charge state	3
Scan Event 1: Full MS	
Microscans (average spectra per analytical scan)	1
Resolution	70 000
Automatic gain control (AGC) target (number of charges to inject and analyse)	5e6
Maximum injection time (IT)	250 ms
Number of scan ranges	1
Scan range	380 - 1500 m/z
Scan Event 2: dd-MS² / dd-Selected Ion Monitoring (SIM)	
Microscans	1
Resolution	17 500
AGC target	5e4
Maximum IT	150 ms
Loop count (repetitions of SE2 per cycle)	20
Spectral multiplexing (MSX) count (number of precursors to be multiplexed)	1
TopN (maximum number of most abundant precursors)	20
Isolation window (isolation width for parent ions)	1.2 m/z
Fixed first mass	100.0 m/z
Normalized collision energy (NCE)	32.5
Stepped NCE	10.0 %
dd Settings (criteria for subsequent ion scans)	
Underfill ratio (minimum percentage of AGC target)	1%
Intensity threshold	3.3e3
Apex trigger (defer acquisition until near peak apex)	- (0)
Charge exclusion (reject individual or undetermined charge states)	Unassigned, 1, 6 to 8, >8
Peptide match	Preferred
Exclude isotopes	On
Dynamic exclusion (prevent ion triggering a subsequent dd-scan and allows analysis of lower abundant ions)	60 s
If idle (select the behaviour following dd-scans)	Do not pick others (default)

4.1.3.6 Proteomic Bioinformatics

Data analysis (Figure 4.4 and Table 4.8) included conversion of raw data files (.RAW), using ProteoWizard MSConvertGUI, to mascot generic format (MGF) files which provide peak list information (Kessner D, Chambers M *et al.*, 2008). Filter parameters for conversion included peak picking and a threshold count of 150. MGF peak lists were searched against a UniProtKB/Swiss-Prot human database (*Homo sapiens*, Canonical sequence, January 2016, Sequences: 20 194) using SearchGUI version 2.3.1 with X!Tandem, MS-GF+ and Comet search engines (Vaudel M, Barsnes H *et al.*, 2011). Post-processing of peptide-spectrum matches (PSMs) for protein identification was done using Peptide Shaker version 1.7.3 (Vaudel M, Burkhart JM *et al.*, 2015).

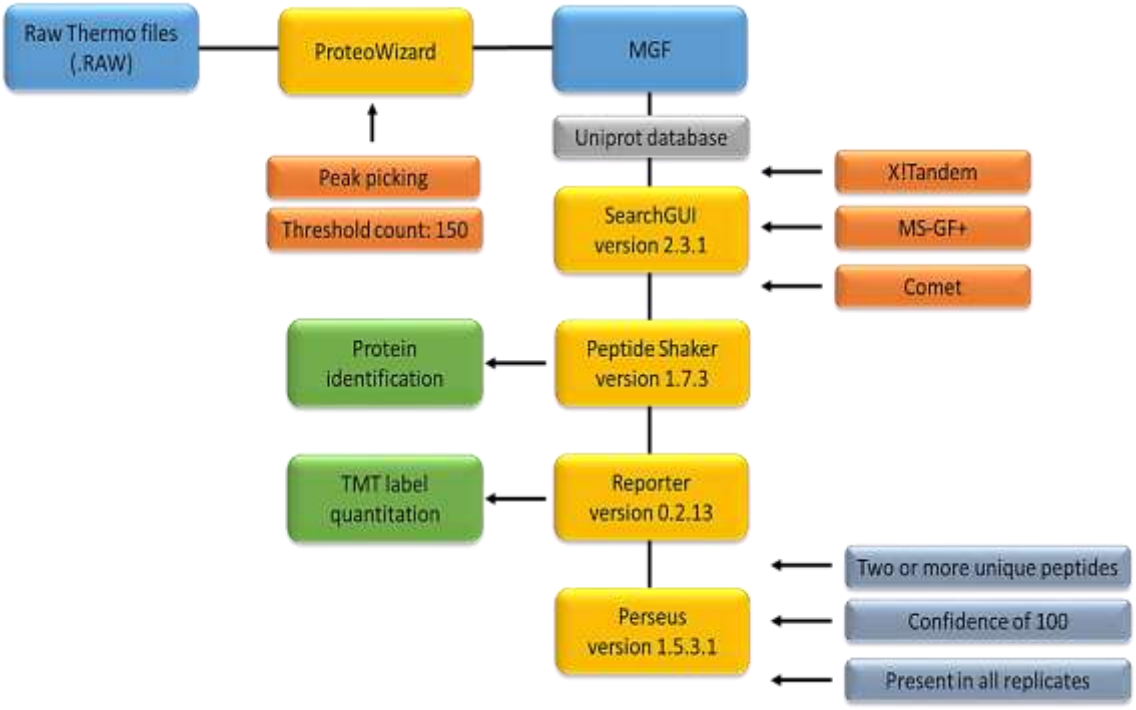

Figure 4.4: Overview of the sequence of data analysis and software packages used.

Peptide Shaker results were imported into Reporter for TMT label quantification (version 0.2.13, *http://compomics.github.io/projects/reporter.html*). Default parameters for Reporter were as follows; intensities were extracted from spectra and deisotoped using label specific purity coefficients. PSM level ratios were estimated using the median of all non-null intensities and normalized. PSM ratios were then aggregated into peptide ratios using a maximum likelihood estimator. Peptide ratios were normalized using the median and aggregated in the same way as the PSMs based on the protein grouping from Peptide Shaker. Protein ratios were normalized similarly to that of peptide ratios.

Results were exported from Peptide Shaker and Reporter and filtered to remove proteins with less than 2 unique peptides and identification confidence of less than 100%. Once filtered, these proteins were used to compare the protein overlap between replicates with global summaries of protein identification, number of peptides, number of PSMs, coverage range and molecular weight range being reported.

Table 4.8: General SearchGUI and Peptide Shaker settings for protein identification

Precursor selection	MS1 precursor
Lowest charge state	2+
Highest charge state	5+
Minimum precursor mass	300 Da
Maximum precursor mass	9000 Da
Minimum collision energy	0 eV
Maximum collision energy	1000 eV
Protein database	SwissProt_Human_January 2016
Decoy database false discovery rate (FDR)	0.01
Enzyme name	Trypsin
Maximum missed cleavage sites	2
Instrument	ESI-ORBITRAP-HCD
Precursor mass tolerance	10 ppm
Fragment mass tolerance	0.2 Da
Fixed Modifications	Carbamidomethyl (C) TMT6plex (K) TMT6plex (N-terminus)
Variable Modifications	Oxidation (M) Deamidated (NQ)
Search engines	X!Tandem MS-GF+ Comet

Quantified and filtered replicate data sets were imported into Perseus (version 1.5.3.1, Max Planck Institute of Biochemistry). For analysis, all NaN (not a number) values were assigned from a normal distribution or as zero to determine the effects of both imputations. Annotations for GOBP (biological processes), GOMF (molecular functions) and GOCC (cellular compartments) were added. Data was generated as histograms for comparisons between replicates, volcano plots, protein profiles and hierarchical clusters. Protein ratios, using averages across replicates, were expressed relative to the most relevant controls within each data set. Drawing valid conclusions from large data sets is tightly coupled to the false-discovery rate (FDR) and appropriate multi-testing corrections to adjust statistical confidence. Multi-sample testing was conducted using ANOVA with a permutation-based FDR for truncation, an FDR of 0.01 with 250 randomizations and no weighting of the mean (s = 0). Volcano plots were generated using two-sided t-tests with an FDR of 0.01 with 250 randomizations with and without weighting of the mean (s = 0 and s = 0.1).

Protein profiles were generated using averages across replicates, where applicable, and protein clusters identified using Euclidean distance from manually produced or curated reference profiles. Where hierarchical clustering was done, row and column trees were pre-processed with k-mean and the average used for linkages. Column clustering was limited by the number of samples and row clustering was limited to 10 - 20 clusters. Principal component analysis (PCA), was done using 5 enrichment components and a Benjamini-Hochberg cut-off method with a 0.05 FDR.

4.1.4 Method Overview

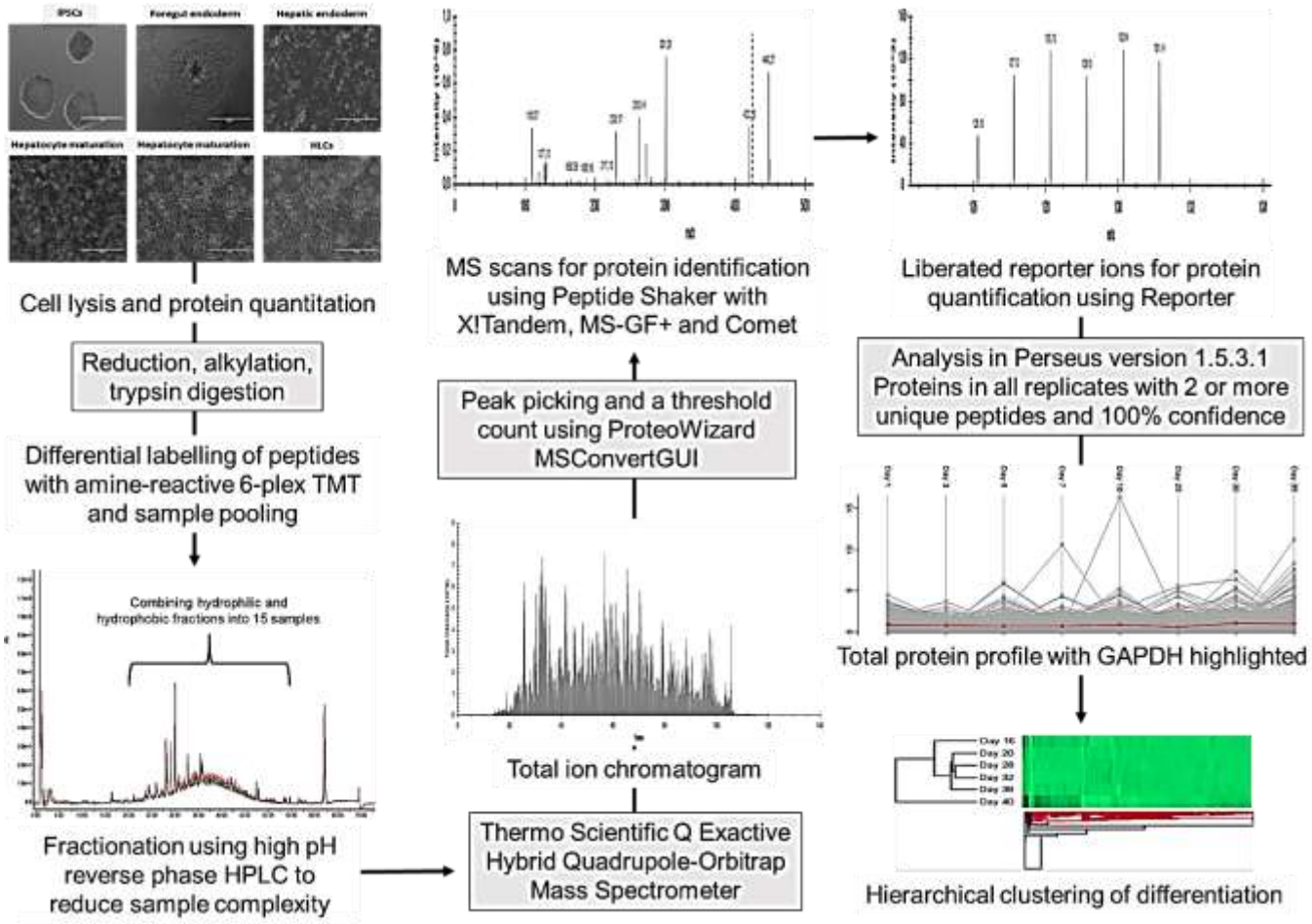

Figure 4.5: The workflow from sample collection to analysis. Whole cell lysates were collected throughout differentiation. Following protein quantitation, 50 µg per sample was reduced, alkylated and digested with trypsin. Peptides derived from each sample were differentially labelled with TMT 6-plex reagent and separated into fifteen one minute fractions using reverse phase high performance liquid chromatography (HPLC). Samples were analysed using liquid chromatography tandem-mass spectrometry (LC-MS/MS) with a linear gradient (4 - 40%) over 120 minute total run time. PSMs were converted to MGFs with peak picking the threshold count filters. Peak lists were searched against a UniProtKB/Swiss-Prot human database using SearchGUI version 2.3.1 with X!Tandem, MS-GF+ and Comet search engines. Output used for further identification and quantification was done using Peptide Shaker version 1.7.3 and Reporter version 0.2.13 respectively. Downstream analysis was conducted in Perseus to generate protein profiles and conduct hierarchical clustering.

4.2 Results and Discussion

Prior to MS-based proteomics, gel based protein-mass profiles of samples were generated to confirm sufficient protein extraction and visualise global protein expression and compare variability in protein expression. (Figure 4.6 and 4.7). As multiple proteins constitute a single band on a gel, the intention here was to highlight general protein band trends rather than make direct comparisons.

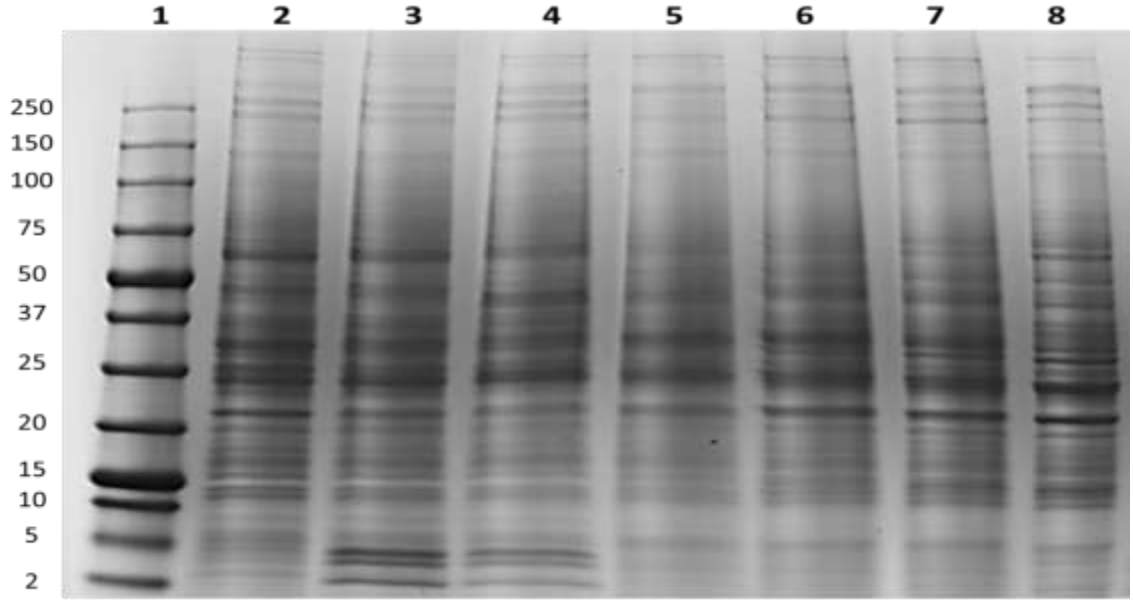

Figure 4.6: Coomassie stained 4-15% Mini-PROTEAN TGX gel of HLC differentiation. 1) Precision Plus Protein Dual Colour standard (2 - 250 kDa), 2) Day 1 induced pluripotent stem cells, 3) Day 3 differentiated endoderm, 4) Day 7 hepatic progenitors and 5 - 8) Hepatic maturation at Days 18, 22, 26 and 32.

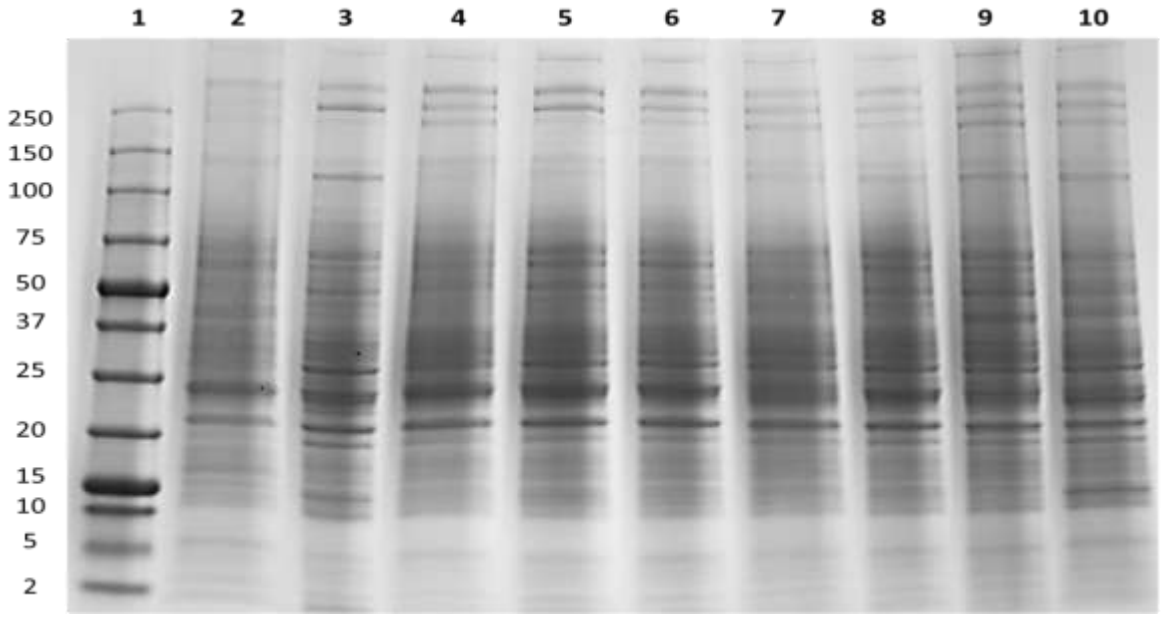

Figure 4.7: Coomassie stained 4-15% Mini-PROTEAN TGX gel of HLCs from different differentiations. 1) Precision Plus Protein Dual Colour standard (2 - 250 kDa) and 2 - 10) HLCs collected at Day 35 from nine differentiations.

The gel-based proteins profiles showed no marked differences but confirmed efficient protein extraction. Samples of equivalent protein were then TMT labelled and mass spectrometry analysis was performed. Sample clean-up and fractionation of each grouped sample improved the quality of the data collected. Pooling of initial UPLC fractions resulting in good orthogonal distribution across total ion chromatograms (TICs) and reliable reporter ion data was evident in the m/z range from 126 - 131 in the MS2 spectra (Figure 4.8D). The raw MS files from each peptide fraction were inspected as TICs, MS1 and MS2 spectra.

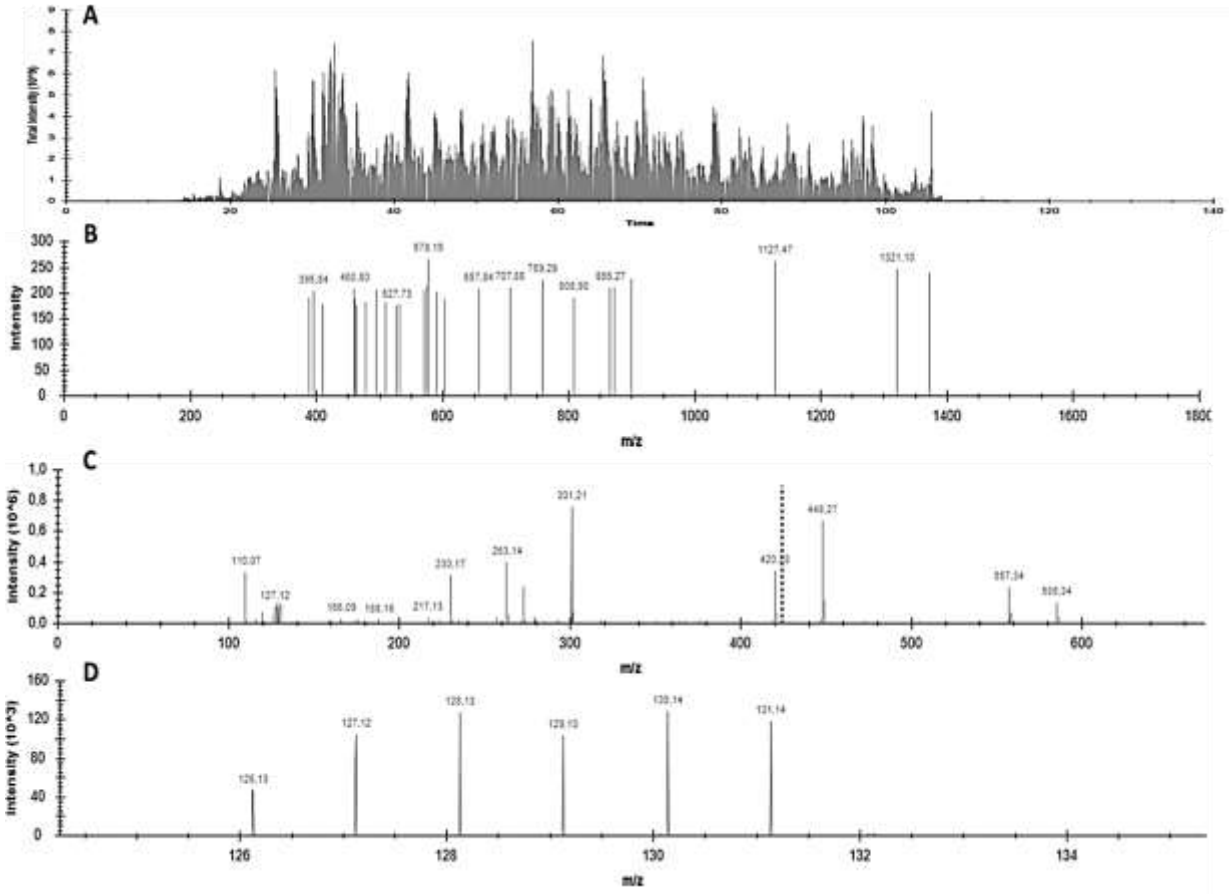

Figure 4.8: LC-MS/MS chromatograms for HLCTC replicate 1. A) Total ion chromatogram over the 120 minute run time, B) MS1 chromatogram, C) MS2 chromatogram and D) MS2 chromatogram zoomed in to illustrate the six reporter ions ranging from 126 - 131.

Factors such as experimental conditions, search engines and bioinformatics approaches should be considered when comparing protein identification across replicates. Protein precipitation can result in the intangible loss of specific subsets of proteins. Additionally, "*under-sampling*" can occur if peptide complexity exceeds chromatographic and mass spectrometer analytical capacity, reducing detection of low-abundance peptides. Low or high stringency in filters can then also contribute to noise, reduced reproducibility or diminished protein identification (Vaudel M, Burkhart JM *et al.*, 2011).

Protein identification for replicates 1 and 2 of the complete differentiation time course (HLCTC) was 6 789 and 6 980 respectively with an overlap of 5 727 proteins once protein groups were filtered. Additionally, protein identification for replicates 1 and 2 of the maturation time-course (HLCLTC) was 6 695 and 6 255 for with an overlap of 5 598 proteins post-filtering (Table 4.9).

Table 4.9: Global summary for replicates of HLC differentiation time courses

	Differentiation time course (HLCTC)		Hepatocyte maturation time course (HLCLTC)	
	Replicate 1	**Replicate 2**	**Replicate 1**	**Replicate 2**
Proteins	6 789	6 980	6 695	6 255
Protein group(s): 2+ unique peptides	6 634	6 792	6 532	6 116
Protein group(s): 2+ unique peptides, high confidence (100%)	6 510	6 672	6 404	6 005
	Protein group(s): 2+ unique peptides, high confidence			
Range of peptide numbers	2 - 387	2 - 372	2 - 423	2 - 419
Range of peptide spectrum matches	2 - 872	2 - 908	2 - 703	2 - 721
Molecular weight range (kDa)	5.02 - 860.12	5.04 - 860.12	5.1 - 1055.3	5.04 - 1010.45
Coverage range	0.18 - 89.87	0.29 - 92.5	0.25 – 92.5	0.4 - 91.63
Proteins in replicate 1 and 2	5727		5598	
Proteins in replicate 1 only	783	-	806	-
Protein in replicate 2 only	-	945	-	407

High protein numbers were identified with a substantial overlap using the methods described. Despite this, the likelihood of detecting proteins which are of low abundance remains a technical limitation of the instrumentation and a feature incorporated by the sample workup. Mass spectrometry acquisition was done using data dependent scans of the top 20 most abundant ions with a dynamic exclusion of 60 seconds. This technical aspect of instrument decision making could bias identification.

In addition, based on the total ion chromatograms (Figure 4.8A), there was no baseline separation despite UPLC fractionation and pooling of hydrophobic and hydrophilic peptides. This suggests a high potential for co-fractionation and co-fragmentation of peptides which limits the detection of low abundance proteins and quantitation accuracy. The percentage of MS/MS spectra assigned could provide an indication of the quality of the data with lower percentages being assigned when peptides co-elute and fragment. However, filtering and threshold counts introduced prior to peptide-spectrum matching means that this is not a definitive analysis strategy. In addition, possible bias could have been incorporated in protein identification based on the fractionation and recombination strategies. This needs to be considered when inferring conclusions based on the mass spectrometry acquisition.

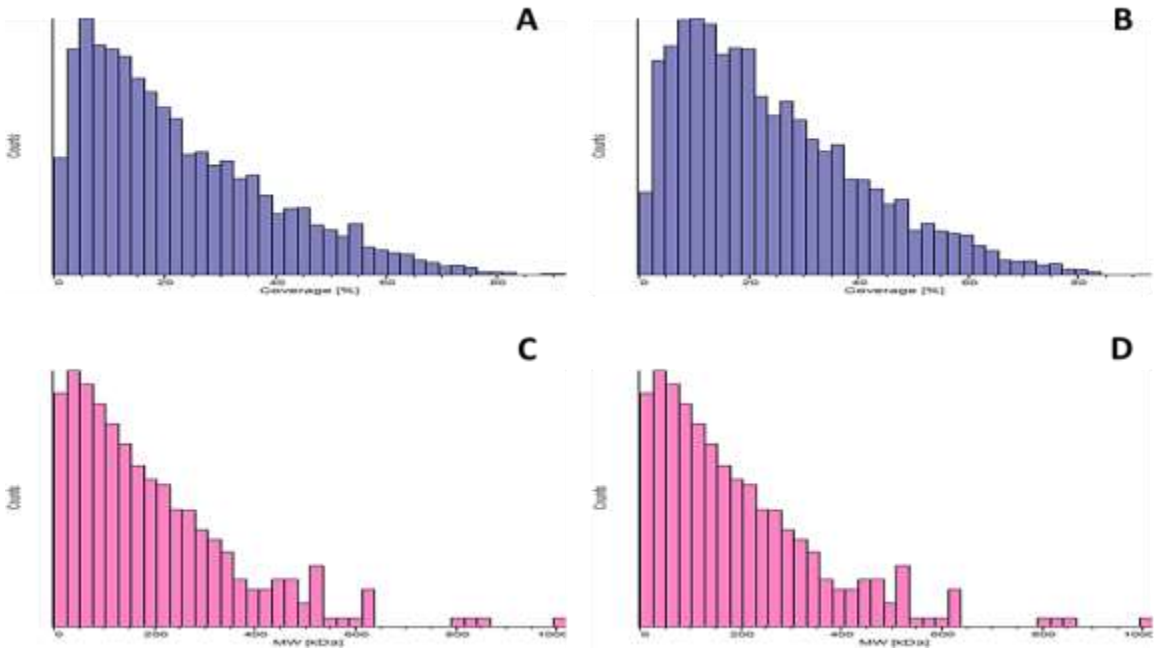

Figure 4.9: Histograms illustrating, A - B) Protein coverage (%) in HLCLTC replicate 1 to 2 and C - D) Molecular weight (kDa) of proteins in HLCLTC replicate 1 to 2 (x-axis: counts).

Histograms of coverage, molecular weight, number of peptides and PSMs showed a high degree of similarity across all time course replicates (Figure 4.9 and 4.10). This combined with the high overlap in number of proteins identified within replicates is suggestive of quality acquisition and data processing. Summary statistics of tandem mass tag distribution following normalization in Reporter (Figure 4.11) were assessed. Despite normalization of quantification data in Reporter, TMT label 128 (HLCTC Day 10) in replicate 1 and TMT label 131 (HLCTC Day 30) in replicate 2 displayed concerning distribution patterns. While having similar median values (0.98 – 1.02) these tags had noteworthy variances in the mean and sum in comparison to other tags. The mean for (HLCTC Day 10) in replicate 1 and TMT label 131 (HLCTC Day 30) in replicate 2 were 0.88 and 1.28 respectively will other tags had a mean within a small variance around 1.00. In order to minimise potential impact of this anomalous distribution on biological findings, these two tags were removed for subsequent analysis.

HLCTC time points were selected to have duplicates of iPSCs at Day 1, anterior definitive endoderm at Day 5, hepatic endoderm at Day 10 and maturing hepatocyte-like cells at Day 30 but also to maximise cover throughout the differentiation with single time points at Day 3, Day 7, Day 25 and Day 35. Despite the intention for duplicate time points throughout the time course, tags with abnormal distribution patterns (HLCTC at Day 10 and Day 30) which were duplicates had to be excluded for further quantitative analysis. While valuable due to coverage of the whole differentiation protocol, proteomic data from the HLCTC needed to be interpreted cautiously as limited replicates were available to account for biological variance associated with differentiation.

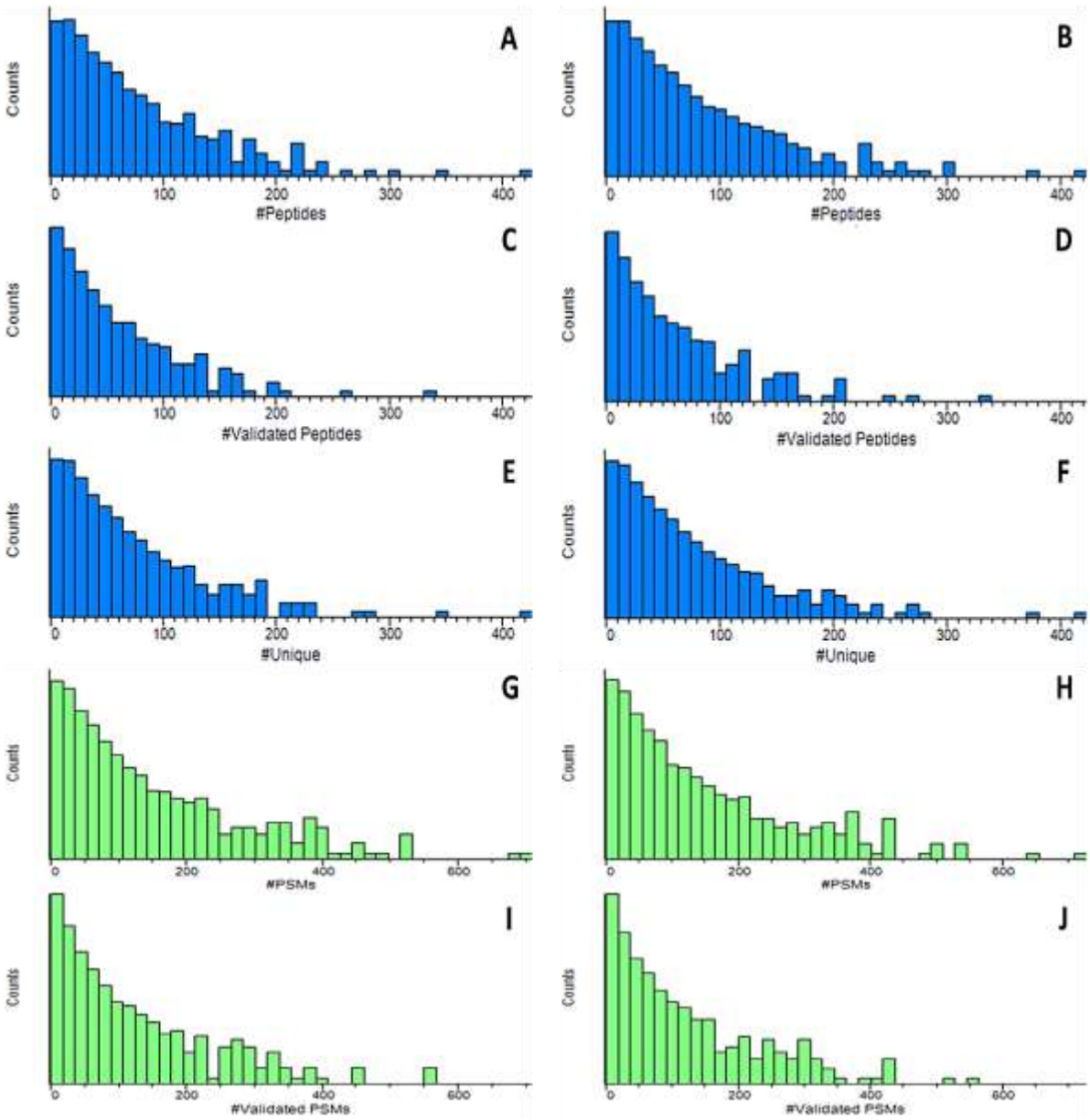

Figure 4.10: Histograms of HLCLTC replicates of A - B) Number of peptides, C - D) Number of validated peptides, E - F) Number of unique peptides, G - H) Number of peptide-spectrum matches and I - J) Number of validated peptide-spectrum matches (x-axis: counts).

However, the late time-course, with the sole focus of hepatocyte maturation (HLCLTC), had no abnormal tag distribution with biological replicates at Day 16, Day 20, Day 28, Day 32 and Day 36 with single time points as Day 24 and Day 40. Protein profiles and numerical averages for proteins identified in both replicates of the time course were used to investigate various features of the HLC proteome.

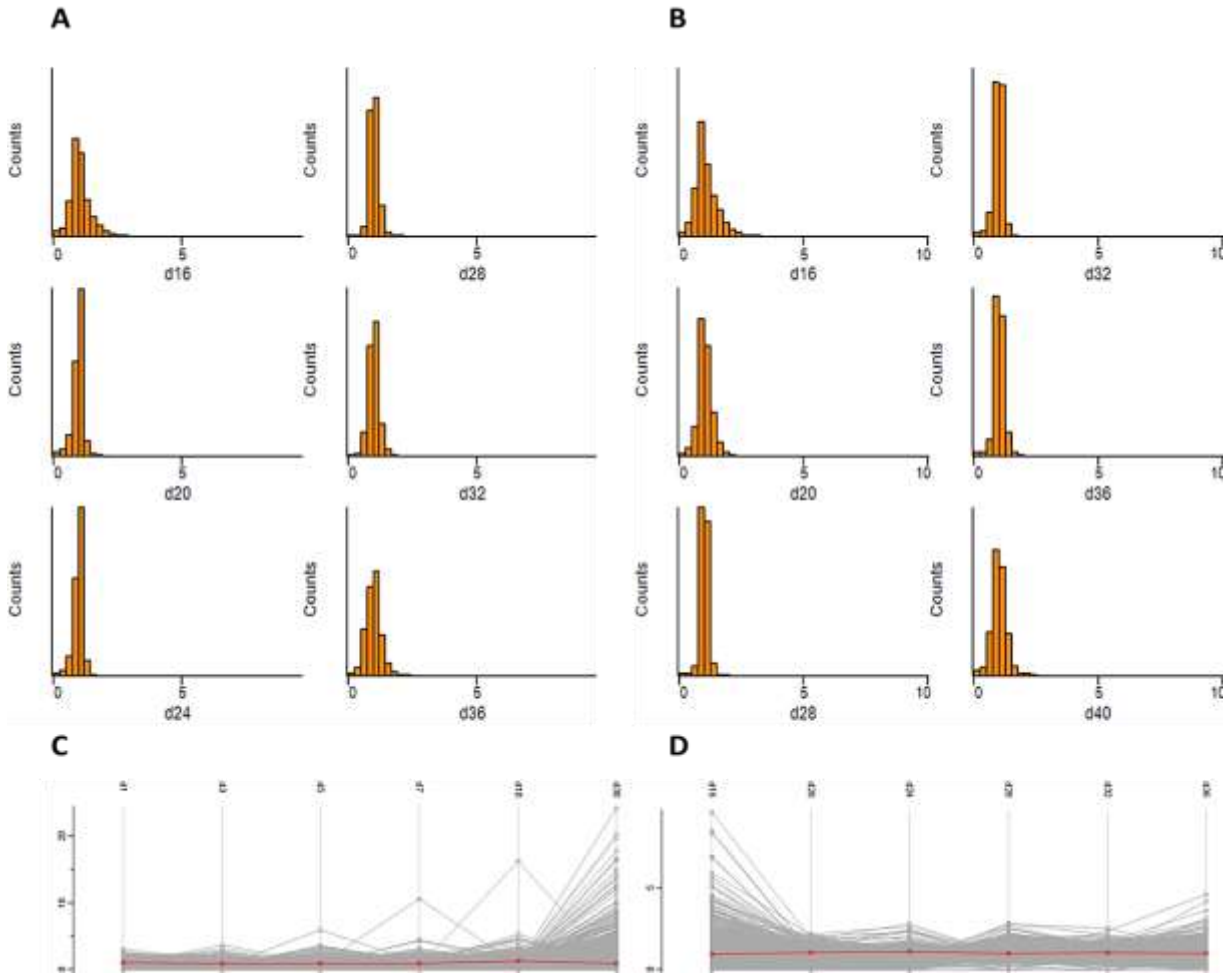

Figure 4.11: A) Histogram of TMT distribution profiles for HLCLTC replicate 1, B) Histogram of TMT distribution profiles for HLCLTC replicate 2, C) Average protein profile for HLCTC and D) Average profile for HLCLTC. Red line on protein profile: actin cytoplasmic 1 (P60709).

4.2.1 Protein Markers and Protein Profiles for HLCTC

Embryonically, the liver develops from the endoderm. Morphogenic movements aid in foregut endoderm patterning which delineates domains for organs including the thyroid gland, stomach, lungs, pancreas and liver. Following endodermal patterning, the liver is derived from the ventral wall of the foregut endoderm consisting mainly of bipotent hepatoblasts, the liver precursor cells. Differentiation of hepatoblasts into hepatic or biliary lineages followed by maturation completes the developmental progression (Lemaigre F, 2010). The highly co-ordinated developmental pattern provides a guideline for hepatocyte differentiation. The aim of conducting a whole proteome time course through HLC differentiation was to determine the *in vitro* correlation with embryonic development.

Cumulatively data suggests the existence of marker sets that can be used to phenotype differentiation status. However, many genes and their respective proteins are not specific to hepatocytes and should therefore be considered in various combinations. DeLaForest *et al.* generated an mRNA fingerprint outlining the various differentiation phases. Genes controlling hepatic phenotypic characteristics and gene expression included FOXA2, TBX3, HHEX, HNF4α, GATA4 and GATA6.

Additionally, gene ontology analysis has revealed that genes associated with hepatic biological processes such as lipid and carbohydrate metabolism, are expressed. However, a limitation of these fingerprints is the variance introduced solely based on protocol duration and specification. Importantly, unique gene sets and protein profiles will continue to evolve as differentiation protocols are optimised (DeLaForest A, Nagaoka M *et al.*, 2011).

The transcriptome and proteome have been investigated to identify markers which distinguish phases of differentiation from iPSCs to hepatocyte-like cells. These include markers which are associated with foetal or adult hepatocytes. Data by Schwartz *et al.* (Table 4.10) provided a basis for interrogation of hepatic marker expression with inclusion of proteins described by many others using various protocols (Agarwal S, Holton KL *et al.*, 2008; Baxter M, Withey S *et al.*, 2015; Duret C, Gerbal-Chaloin S *et al.*, 2007; Lee KD, Kuo TKC *et al.*, 2004; Schwartz R, Fleming H *et al.*, 2014; Si-Tayeb K, Noto FK *et al.*, 2010; Touboul T, Hannan NR *et al.*, 2010).

It is important to note that whole proteome analysis can fail to detect low abundant proteins which are detected at the genomic or transcriptome level due to the ability to amplify DNA and RNA. In addition, sample preparation and the technical aspects of analysis can bias the detection of more hydrophilic, easily ionisable and abundant proteins. These factors need to be considered when defining the correlation to other “omics” data sets.

Assessing the typical molecular markers which govern each phase of differentiation from pluripotency to hepatocytes showed a notable absences of specific proteins. None of the characteristic markers of pluripotency were detected. However, anterior definitive endoderm markers Cerberus (O95813) and FGF-17 (O60258) were identified in HLCTC replicate 1. Markers for hepatic endoderm and hepatoblasts were found with reasonable consistency across replicates. Importantly, a large number of adult and foetal hepatocyte markers were identified and analysed in this study.

Table 4.10: Marker expression throughout hepatocyte differentiation
(Schwartz R, Fleming H *et al.*, 2014)

Differentiation phase	Molecular marker
Pluripotency	NANOG, OCT-4, SOX-2, SSEA-3/4, TRA 1-60
Anterior definitive endoderm	*Cerberus*, CXCR-4, *FGF-17*, *FOXA2*, Goosecoid, HHEX, MixL1, SOX-17, GATA-4
Hepatic endoderm	***APOA1***, ***APOB***, BMP-6, *Cerberus*, CXCR-4, DUSP6, *FOXA2*, GATA-4, HHEX, ***HNF4α***, SOX-17 (low), *TBX-3*
Hepatoblast	***APOA1***, ***APOB***, ***AFP***, ***ALB***, Decorin, FOX-A2, GATA-4, GSTA1, HNF1α, ***HN-4α***, HHEX, KRT19, **SOX-17** (low), *TBX-3*, ***TTR***
Adult hepatocyte	***A1AT***, ***ALB***, ***AFP***, ***APOA1***, ***APOB***, APOF, ***ASGR1***, BSEP, ***CK8***, ***CK18***, CPS1, *CYP1A2*, *CYP2A6*, *CYP2B6*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A4*, CYP3A7 (low), Decorin, FOX-A2, ***GATA-4***, *GSTA1*, HNF1α, ***HNF4α***, HHEX, ***MAOA***, ***MAOB***, ***MRP2***, **SOX-17** (low), TBX-3, ***TTR***, ***UGT1A1***
Foetal hepatocyte	***A1AT***, ***APOA1***, ***APOB***, ***AFP***, ***CK8***, ***CK18***, ***CYP3A7***, Decorin, FOX-A2, ***GATA-4***, *GSTA1*, HNF-1α, ***HNF4α***, HHEX, ***KRT19***, **SOX17** (low), TBX-3, **TTR**

* Proteins in bold were identified in both replicates of at least one time course
Proteins underlined were identified in at least one replicate of either time course

Molecular markers of hepatic phenotypes mostly peaked on Day 30 of differentiation with only 3 of the 17 proteins identified not elevated at Day 30 (Table 4.11). This suggests that if these proteins are true markers of hepatic maturity, whether adult or foetal, that differentiating beyond 30 days is potentially associated with a reduction in the hepatic phenotype. In addition, the major metabolizing cytochromes, which represent an adult hepatic phenotype, were only identified in one replicate (HLCTC replicate 2). This has numerous implications for the conclusions made from this data set. It cannot be established whether this is due to loss of hydrophobic peptides during preparative steps, acquisition bias based on protein abundance between specific replicates or merely the absence of the protein in the sample. Detection of these proteins in only one replicate suggests that there is variability throughout differentiation in addition to assay variability.

As only one biological replicate was included for each of Day 30 and Day 35, these results could be considered an artefact of a specific differentiation. Despite this, the trend from low expression of many hepatic proteins in iPSC, endoderm and hepatic endoderm to higher expression in hepatocyte-like cells illustrates successful augmentation of the proteome which mimics the morphological observations seen throughout differentiation.

Table 4.11: Pseudo-heatmap of proteins present in the HLCTC which are suggestive of hepatic phenotypes expressed as a ratio relative to iPSCs

Protein description	iPSC	Endoderm	Foregut endoderm		Hepatic endoderm	Hepatocyte-like cells			Main accession
	Day 1	Day 3	Day 5	Day 7	Day 10	Day 25	Day 30	Day 35	
Albumin	1.000	0.845	0.876	4.162	2.617	1.651	8.823	1.546	P02768
Alpha-1-antitrypsin	1.000	1.382	0.913	0.000	0.000	2.300	3.808	2.584	P01009
Alpha-2-HS-glycoprotein	1.000	1.316	0.720	1.245	1.590	1.892	1.667	1.628	P02765
Alpha-fetoprotein	1.000	0.813	0.516	1.050	1.160	7.520	6.275	2.098	P02771
Amine oxidase [flavin-containing] A	1.000	1.355	0.779	1.123	1.107	1.796	3.016	0.926	P21397
Amine oxidase [flavin-containing] B	1.000	1.162	0.932	1.087	1.133	1.536	2.464	1.063	P27338
Apolipoprotein	1.000	1.411	0.355	1.018	2.032	4.575	2.814	2.900	P02647
Apolipoprotein	1.000	1.237	0.759	0.819	1.263	2.271	2.927	1.186	P04114
Carbamoyl-phosphate synthase	1.000	1.343	0.999	1.469	1.138	1.754	8.899	1.701	P31327
Fibrinogen gamma chain	1.000	1.581	0.472	0.000	0.886	2.562	3.445	2.327	P02679
Fibrinogen alpha chain	1.000	3.497	1.549	0.000	5.097	7.475	9.368	6.782	P02671
Fructose-bisphosphate aldolase B	1.000	1.229	0.458	0.720	0.926	0.828	3.492	1.374	P05062
Glutathione peroxidase 1	1.000	0.861	0.778	0.983	0.871	0.462	0.739	0.758	P07203
Keratin. type I cytoskeletal 18	1.000	0.962	1.019	1.177	1.922	1.532	1.279	1.373	P05783
Keratin. type II cytoskeletal 8	1.000	0.918	1.069	1.252	2.083	2.128	1.470	1.420	P05787
Transthyretin	1.000	2.587	0.226	2.754	2.103	8.942	7.768	2.053	P02766
UDP-glucuronosyl transferase	1.000	1.599	1.042	1.657	1.825	1.625	4.427	7.476	P22309

Si-Tayeb *et al.* identified 40 mRNAs which, to varying degrees, when derived from iPSCs, are solely expressed in the liver. Within the HLCTC data set, only eight proteins correlating to these mRNAs were identified and quantified. Serine-pyruvate aminotransferase (P21549), haptoglobin (P00738), s-adenosylmethionine synthase isoform type-1 (Q00266), prothrombin (P00734), angiotensinogen (P01019), heparin cofactor 2 (P05546) and alpha-2-HS-glycoprotein (P02765) were elevated at Day 30 (Si-Tayeb K, Noto FK *et al.*, 2010).

The notable variance in the relative ratio of some proteins, despite undergoing the same and seemingly adequate normalisation, suggests that differentiation is not as tightly regulated as would be required for use in safety pharmacology and toxicology. An advantageous strategy may have been to pool multiple differentiations per tag to include more differentiations (50 µg from two or three differentiations at the same time point) may have provided a more accurate representation of relative abundance.

In addition to observations of known markers throughout differentiation, reference protein profiles (Figure 4.12) were manually generated to discover proteins which peaked at a specific time point but maintained linearity throughout the rest of the time course. Using these profiles, the top 10 - 20 proteins clustering together were identified and annotated with the molecular or biological functions (Table 4.12) to assess correlation of protein function with sequential phases of differentiation.

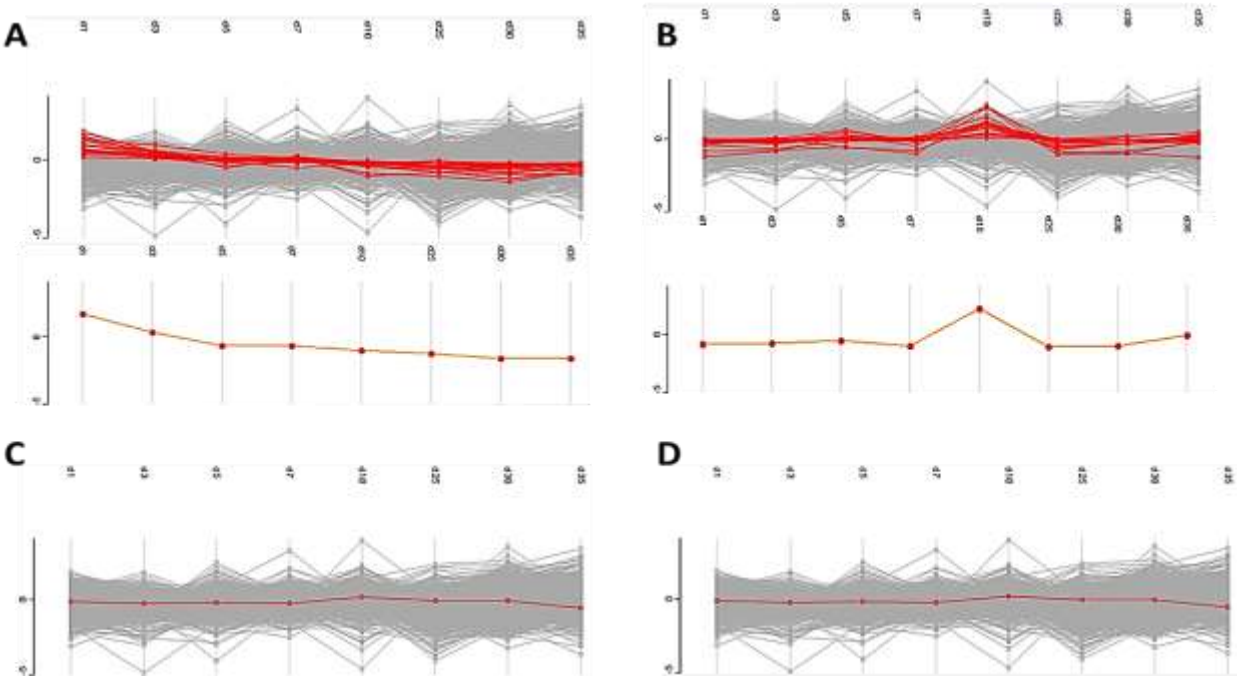

Figure 4.12: Protein profiles (top) and reference profiles (bottom), from Day 1 to Day 35, generated to cluster the top 20 proteins which peak in A) iPSCs on Day 1, B) Hepatic endoderm on Day 10. C) Protein profile illustrating actin cytoplasmic 1 (P60709) and D) Protein profile of tubulin beta-6 chain (Q9BUF5).

Most of the distinct protein peaks (Table 4.12) corresponded to functions required during each differentiation phase. Of interest were those associated with cell cycle regulation such as telomeric repeat-binding factor-1 (P54274) in iPSCs and leucine zipper putative tumour suppressor-1 (Q9Y250) during foregut endoderm associated with proliferation. Additionally, expression of neural cell adhesion molecule-1 (P13591), which is responsible for the organogenesis of mesodermal and endodermal derivatives peaked at Day 5 during foregut endoderm formation.

In hepatic endoderm, proteins involved in multicellular organismal development such as protein shisa-2 homolog (Q6UWI4) and vang-like protein 1 (Q8TAA9) were prominent whereas HLCs express high levels of cadherin-17 (Q12864), a protein involved in morphological organization of liver (TheUniProtConsortium, 2014). Within the co-ordinated developmental pattern of embryonic development, expression of proteins including SRY-box containing transcription factor 17 (SOX17, Q9H6I2) and chemokine CXC receptor 4 (CXCR4, P61073), govern development of regionalized endoderm (Lemaigre F, 2010). In this data set, SOX17 which modulates differentiation via WNT3A, peaked at Day 5 during foregut endoderm differentiation. Despite CXCR4 not being identified in this data set, the presence of SOX17 suggests the initial pattern of embryonic development was followed.

Several signalling factors then confer hepatic identity of hepatoblasts which is evidenced using the earliest liver markers albumin, transthyretin, and α-fetoprotein. These proteins were present and increased over the time course constituting a reliable way of confirming the hepatocyte differentiation pathway. Additionally, transcription factors required for hepatic development but not unique to the liver, including HNF3α/FOXA1, HNF3β/FOXA2, and HNF1β/TCF were identified but not in all of the replicates (Lemaigre F, 2010).

Hepatic lineages then mature to acquire metabolic properties under a network of six core transcription factors. These factors, whose activity is governed by their abundance, are HNF1α, HNF1β, HNF3β, HNF4α1, HNF6 and liver receptor homolog (LRH) 1 which functionally cross-regulate (Lemaigre F, 2010). While not present in all proteomic data sets, these factors suggest that the time course of differentiation (Day 1 - Day 35) mimics that of embryonic development.

Table 4.12: Top proteins which peak at various stages of the protein profile but remain consistent throughout the rest of the differentiation time-course (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in iPSCs (Day 1)		
Synaptic vesicle membrane protein VAT-1 homolog-like	Q9HCJ6	Oxidoreductase activity Zinc-ion binding
Lymphoid-specific helicase	Q9NRZ9	ATP-binding and helicase activity Cell division and mitotic nuclear division Multicellular organismal development
Telomeric repeat-binding factor 1	P54274	DNA, ubiquitin and microtubule binding Cell division, G2/M transition, meiotic telomere clustering
Alpha-aminoadipic semialdehyde synthase, mitochondrial	Q9UDR5	Saccharopine dehydrogenase activity (lysine metabolism) Cellular nitrogen compound metabolic process Catabolic processing to acetyl-CoA
DNA (cytosine-5)-methyltransferase 3A	Q9Y6K1	Chromatin and mitotic cell cycle regulator DNA and metal ion binding and DNA methylation

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in foregut endoderm (Day 5)		
Tyrosine-protein kinase ZAP-70	P43403	ATP binding Protein tyrosine kinase activity Adaptive immune responses
Leucine zipper putative tumor suppressor 1	Q9Y250	Cell cycle regulation Stabilize active CDC2-cyclin B1 complex
Neural cell adhesion molecule 1	P13591	Cell adhesion molecule Expressed during organogenesis of mesodermal and endodermal derivatives
DNA dC->dU-editing enzyme APOBEC-3G	Q9HC16	Cytidine and deoxycytidine deaminase activity which maintains the cellular pyrimidine pool
Collagen alpha-2(V) chain	P05997	Member of group I collagen (fibrillar forming collagen) Extracellular matrix structural constituent Key determinant in the assembly of tissue-specific matrices
Transcription factor SOX-17	Q9H6I2	Modulates transcriptional regulation via WNT3A Normal development of definitive gut endoderm
Proteins peaking in hepatic endoderm (Day 10)		
Protein shisa-2 homolog	Q6UWI4	Developmental protein Multicellular organismal development
Frizzled-5	Q13467	Coupled to the beta-catenin canonical signaling pathway Inhibition of glycogen synthase kinase Nuclear accumulation of β-catenin (cellular adhesion)
LIM and cysteine-rich domains protein 1	Q9NZU5	Transcriptional cofactor that restricts GATA6 function Represses GATA6-mediated transactivation of lung and cardiac tissue-specific promoters
Vang-like protein 1	Q8TAA9	Multicellular organismal development Wnt signaling pathway, planar cell polarity pathway
β-1,3-galactosyl-O-glycosyl-glycoprotein β-1,6-N-acetyl-glucosaminyltransferase	Q02742	Pathway protein glycosylation (protein metabolism) Cell adhesion molecule production
Proteins peaking in hepatocyte-like cells (Day 30)		
Serine--pyruvate aminotransferase	P21549	Transaminase activity (amino acid and α-keto group) Amino acid binding
Cadherin-17	Q12864	Calcium-dependent cell adhesion proteins May have a role in the morphological organization of liver
Sialate O-acetylesterase	Q9HAT2	Sialate O-acetylesterase activity Carbohydrate metabolic process
Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	P21695	Cellular lipid metabolic process Major link between carbohydrate and lipid metabolism
Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial	Q96I99	Catalyzes GTP-dependent ligation to form succinyl-CoA Tricarboxylic acid pathway (carbohydrate metabolism)

Hierarchical clustering was done using averaged group data (Figure 4.13). Generating ten row clusters for individual samples produced clusters ranging from 3 to 5 483 (95.95% of identified proteins) proteins per cluster. Intermediately sized clusters included 34, 41, 67 and 68 proteins. Small groups of 3 to 13 proteins demonstrated no trends while clusters containing 67 and 68 increased over the time-course until Day 30 or 35 (Figure 4.13B and C) respectively. Samples generally co-segregated according to differentiation phases with major nodes distinguishing samples from Day 1 to 10 and Day 25 to 35.

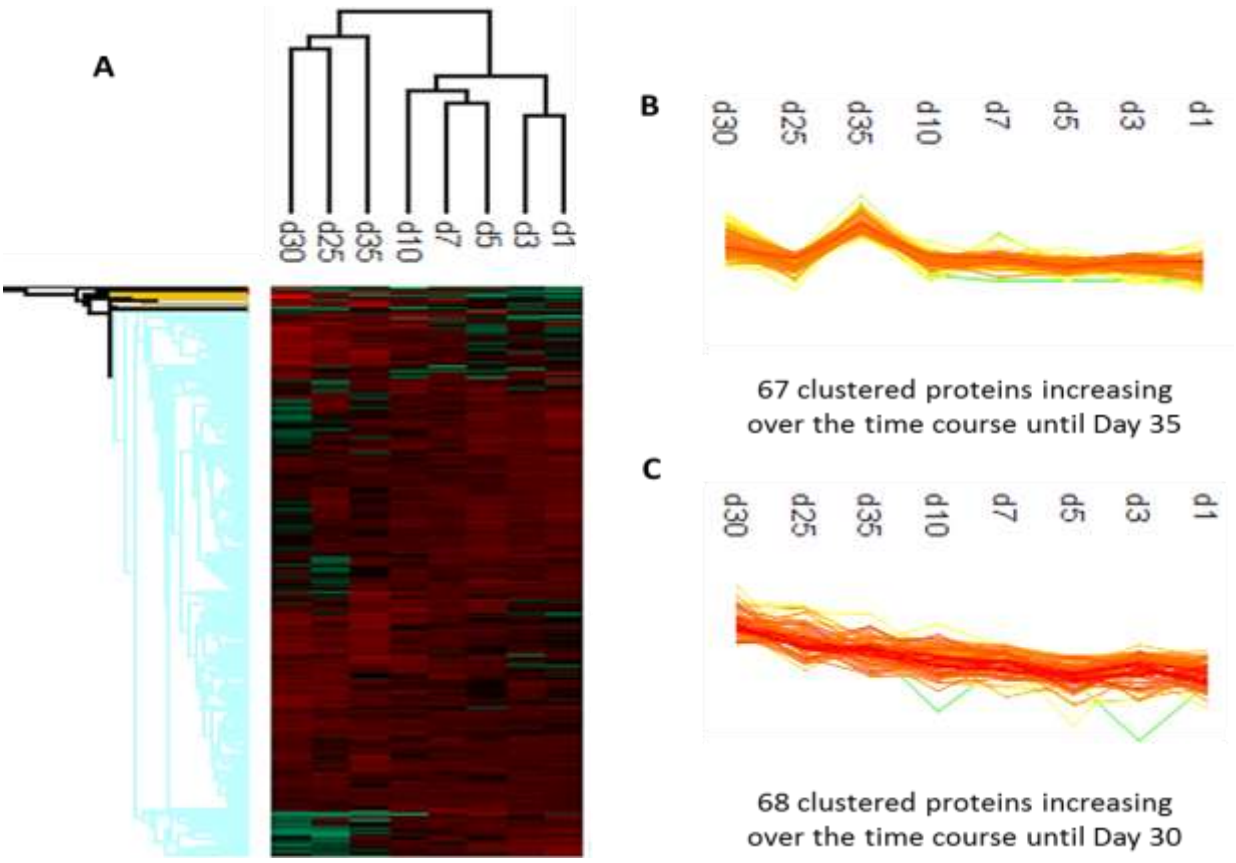

Figure 4.13: A) Hierarchical clustering through the HLC time-course from Day 1 to Day 35. B - C) Protein profiles of clusters demonstrating interesting trends (proteins listed in appendix).

Differentiation of iPSCs to HLCs requires a complex interplay of cell proliferation and maturation. In this differentiation protocol, following foregut endoderm induction (Day 8), cells were reseeded at a density to ensure confluence. Cells should no longer be undergoing active cell division once seeded at confluence and a decrease in proliferation markers should be evident. To observe whether this was evident in the proteome, ten markers of cell growth and proliferation were identified (Table 4.13).

Ki-67 (P46013), responsible for maintaining cell proliferation, cyclin-dependent kinase-1 (P06493), responsible for G2/M cell cycle transition as well as cyclin-dependent kinase 4 (P11802) responsible for G1/S transition gradually declined between Day 1 and Day 5. This was also evident for Geminin (O75496), which when degraded during metaphase-anaphase transitions, permits replication in the subsequent cell cycle and M-phase inducer phosphatase-3 (P30307) which directly dephosphorylates CDK1 and stimulates its kinase. The opposite trend was found for cyclin-dependent kinase inhibitor 1B (P46527), a protein governing G1 arrest (inhibitor of cyclin E- and cyclin A-CDK2 complexes), which increased throughout the differentiation. These proteins illustrate that cell proliferation declined throughout the 35 day differentiation protocol while markers of hepatic phenotypes increased.

Table 4.13: Pseudo-heatmap of proteins involved in cell growth, proliferation or survival which are generally expressed in early differentiation but decrease as hepatocyte-like cells mature expressed as a ratio relative to iPSC

Protein description	iPSC	Endoderm	Foregut endoderm		Hepatic endoderm	Hepatocyte-like cells			Main accession
	Day 1	Day 3	Day 5	Day 7	Day 10	Day 25	Day 30	Day 35	
Antigen KI-67	1.000	0.653	0.558	0.447	0.476	0.000	0.000	0.241	P46013
Cyclin-dependent kinase 1	1.000	1.055	1.112	0.767	0.772	0.000	0.000	0.392	P06493
Cyclin-dependent kinase 4	1.000	1.090	1.015	0.958	0.861	1.002	0.755	0.896	P11802
DNA replication licensing factor MCM3	1.000	0.983	0.601	0.584	0.523	0.072	0.165	0.314	P25205
Endothelial differentiation-related factor 1	1.000	0.837	0.789	0.695	0.715	0.843	0.465	0.577	O60869
Far upstream element-binding protein 3	1.000	1.129	0.968	1.040	1.101	1.550	1.062	1.011	Q96I24
Geminin	1.000	0.988	0.524	0.561	0.445	0.417	1.329	0.407	O75496
M-phase inducer phosphatase 3	1.000	0.523	0.808	0.532	0.427	0.338	0.395	0.452	P30307
Phosphoinositide 3-kinase regulatory subunit 4	1.000	1.141	0.900	0.925	0.834	0.868	0.871	0.996	Q99570
Proliferating cell nuclear antigen	1.000	0.749	0.698	0.716	0.685	0.000	0.000	0.257	P12004
Caspase-3	1.000	1.042	0.891	1.124	0.849	0.992	0.971	0.701	P42574
Cyclin-dependent kinase inhibitor 1B	1.000	0.412	1.807	1.895	3.012	2.354	2.622	2.947	P46527

4.2.2 Protein Markers and Protein Profiles for HLCLTC

Again, using the mRNAs designated by Si-Tayeb *et al.* (2010) which provide a transcriptional fingerprint of the differentiated hepatic phenotype, 17 correlating proteins were expressed in the maturation time course. Proteins included serine-pyruvate aminotransferase (P21549), haptoglobin protein (P00738), complement factor B (P00751), S-adenosylmethionine synthase isoform type-1 (Q00266), antithrombin-III (P01008), protein AMBP (P02760), complement 5 (P01031), prothrombin (P00734), angiotensinogen (P01019), ribonuclease 4 (P34096), heparin cofactor 2 (P05546), beta-2-glycoprotein 1 (P02749), fibrinogen alpha chain (P02671), vitronectin (P04004), fibrinogen gamma chain (P02679), alpha-2-HS-glycoprotein (P02765) and apolipoprotein A-II (P02652).

Identification of 17 proteins here as opposed to 8 in the complete differentiation time-course could be associated with the fact that all cells during maturation should be expressing these proteins or that duplicates per time point were included for analysis.

During the maturation time course, molecular markers of hepatic phenotypes remained relatively consistent with a general peak in cells collected at Day 32 (Table 4.14). This is similar to the complete differentiation time course where abundance was greatest at Day 30. This combination of data suggests that proteins reflecting a hepatic phenotype may decline between Day 30 and 35 and provide a proteomic impression of stagnated differentiation. This notion can only be suggested as this phenomenon could be associated with these specific biological differentiation series.

Despite a large number of hepatic molecular markers being identified and quantified, there was a noteworthy absence of important metabolizing enzymes. None of the major CYP isoforms (CYP1A2, CYP3A4, CYP3A5, CYP2B6, CYP2C9, CYP2C19, and CYP2D6) were identified in either time course. As mentioned previously, this could be of various origins including, protein abundance, differentiation heterogeneity, preparative, instrumental and data processing.

Si-Tayeb *et al.* noted that despite both ES cells and iPSCs producing hepatocyte characteristics with high efficiency, the cohort of mRNAs encoding phase I and phase II enzymes was incomplete (Si-Tayeb K, Noto FK *et al.*, 2010). Culture conditions required to express the full panel of phase I and II enzymes in an artificial environment will continually be improvement as the understanding of developmental biology improves.

Table 4.14: Pseudo-heatmap of proteins present in the HLCLTC which are suggestive of hepatic phenotypes expressed as a ratio relative to HLCs at Day 16

Protein description	Hepatocyte-like cells						Main accession
	Day 16	Day 20	Day 28	Day 32	Day 36	Day 40	
Albumin	1.000	1.465	1.028	1.172	0.903	0.557	P02768
Alpha-1-antitrypsin	1.000	0.766	1.376	1.355	0.996	0.785	P01009
Alpha-2-HS-glycoprotein	1.000	4.905	6.754	5.502	4.146	4.249	P02765
Alpha-fetoprotein	1.000	1.015	1.550	1.567	1.194	1.043	P02771
Amine oxidase [flavin-containing] A	1.000	1.406	1.981	2.023	1.410	0.421	P21397
Amine oxidase [flavin-containing] B	1.000	1.100	1.545	1.696	1.343	0.599	P27338
Apolipoprotein A	1.000	0.849	1.605	1.480	1.122	0.778	P02647
Apolipoprotein B	1.000	0.874	1.250	1.430	0.907	0.828	P04114
Carbamoyl-phosphate synthase	1.000	0.730	0.872	1.087	0.974	0.260	P31327
Fibrinogen alpha chain	1.000	0.816	1.946	0.906	0.940	0.986	P02671
Fibrinogen gamma chain	1.000	0.854	2.077	1.117	1.154	1.051	P02679
Fructose-bisphosphate aldolase B	1.000	0.632	0.620	0.836	0.602	0.282	P05062
Glutathione peroxidase 1	1.000	0.796	0.851	1.031	0.915	0.620	P07203
Keratin. type I cytoskeletal 18	1.000	0.789	1.015	0.842	0.975	1.034	P05783
Keratin. type II cytoskeletal 8	1.000	0.927	1.062	0.892	1.077	1.209	P05787
Transthyretin	1.000	0.927	1.062	0.892	1.077	1.209	P02766
UDP-glucuronosyltransferase 1	1.000	0.604	0.542	0.622	0.468	0.126	P22309

Proteins of interest from protein profiles (Figure 4.14) at Day 32 included those involved in carbohydrate biosynthetic and metabolic processes, cellular lipid metabolic processes and the mitochondrial membrane respiratory chain (Table 4.15). At Day 36, F-box only protein 30 (Q8TB52) and protein cereblon (Q96SW2), which are involved in protein ubiquitination acting as signals for protein degradation by the proteasome, peak. This increase in ubiquitination could be indicative of cellular stress as ubiquitination is also involved in protein relocation and interactions. This is suggestive of a general decline in active metabolic processes and increasing protein degradation which could points to a proteomic window for ceasing differentiation.

A concerning feature of the relative ratios presented here is that few of these proteins demonstrate a notable fold-change that meets a distinct statistical cut-off. This could be due to the fact that the relative ratios are being compared to cells differentiated to Day 16 which have transitioned from hepatic endoderm to HLCs. Applying additional statistics between days throughout the time-course would be beneficial to identify a cohort of proteins which vary significantly over this specific time course. As only two values per day were collected statistical testing was not realistic. Lack of statistical rigor is a concern and the ability to implement other statistic tests is a limitation. However, GO enrichment and further detailed pathway analysis could prove insightful in extracting proteins not typically annotated as hepatic markers which do more robustly characterise this data.

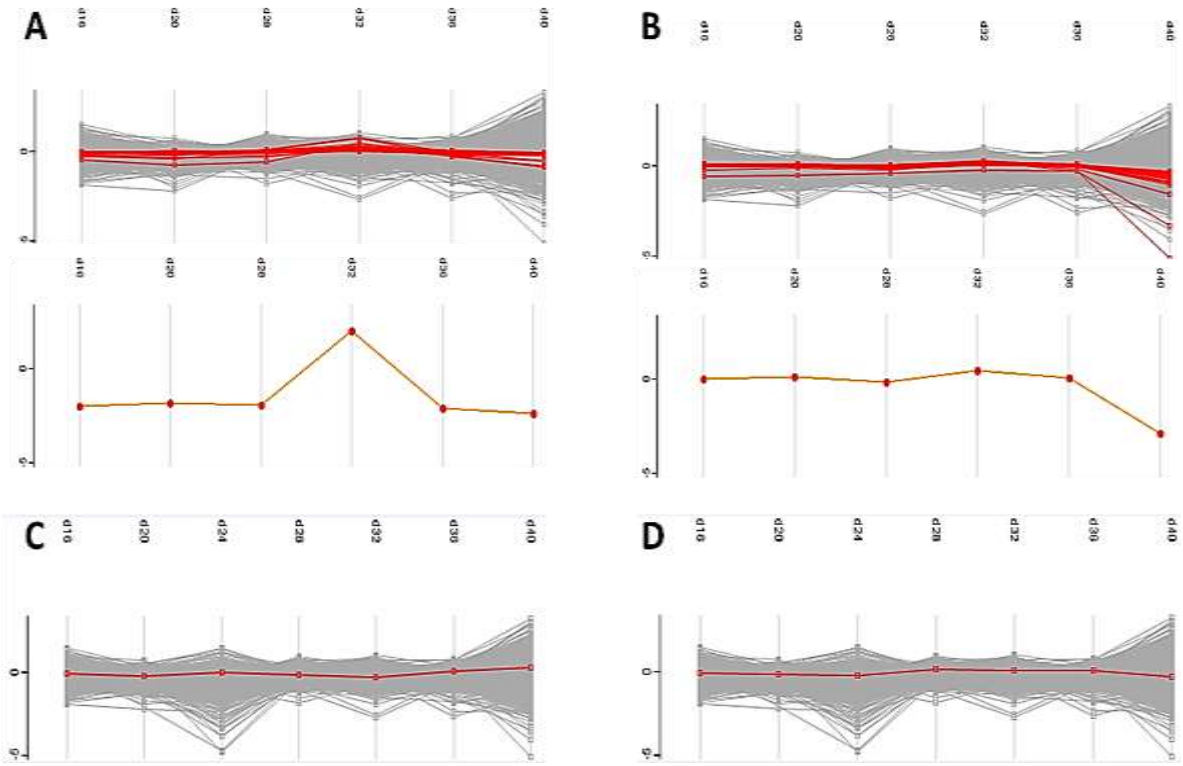

Figure 4.14: Protein profiles (top) and reference profiles (bottom), from Day 16 to Day 40, generated to cluster the top 20 proteins which A) Peak in HLCs on Day 32, B) Decline in HLCs on Day 40. C) Protein profile of actin cytoplasmic 1 (P60709) and D) Protein profile of glyceraldehyde-3-phosphate dehydrogenase (GAPDH; P04406).

Table 4.15: Top proteins which peak at various stages of the protein profile but remain consistent throughout the rest of the differentiation time course (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in hepatocyte-like cells (Day 32)		
Carbohydrate sulfotransferase 13	Q8NET6	N-acetylgalactosamine 4-O-sulfotransferase activity Carbohydrate biosynthetic and metabolic processes Modify glycolipids or glycoproteins along secretory pathways
Acyl-CoA desaturase	O00767	Cellular lipid metabolic process Contributes to the biosynthesis of membrane phospholipids, Cholesterol esters and triglycerides
Solute carrier family 2, facilitated glucose transporter member 5	P22732	Fructose and glucose transmembrane transporter Carbohydrate metabolic process
Cytosolic Fe-S cluster assembly factor NARFL	Q9H6Q4	Component of cytosolic iron-sulfur protein assembly complex Iron-sulfur cluster assembly (protein co-factor)
NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13	Q9P0J0	Accessory subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I) Transfer of electrons from NADH to respiratory chain
Proteins peaking in hepatocyte-like cells (Day 36)		
F-box only protein 30	Q8TB52	SCF-type E3 ubiquitin ligase complex Protein ubiquitination
Myelin basic protein	P02686	Structural constituent of myelin sheath Response to fatty acids and toxic substances
Endoplasmin	P14625	Molecular chaperone that functions in the processing and transport of secreted proteins Cellular protein metabolic process
Protein disulfide-isomerase A4	P13667	Catalyzes the rearrangement of -S-S- bonds in proteins Cell redox homeostasis Chaperone-mediated protein folding
Protein cereblon	Q96SW2	DCX protein ligase complex Protein ubiquitination
Proteins peaking in hepatocyte-like cells (Day 40)		
Excitatory amino acid transporter 1	P43003	Glutamate binding and biosynthetic processes L-glutamate transmembrane transporter activity
DNA polymerase delta catalytic subunit	P28340	DNA synthesis (polymerase) Exonucleolytic activity (degrades single stranded DNA) Damaged DNA binding
Ribonucleases P/MRP protein subunit POP1	Q99575	Component of ribonuclease P Endonucleolytic cleavage of RNA
HMG box transcription factor BBX	Q8WY36	Transcription factor for cell cycle progression (G1 to S phase)
40S ribosomal protein S29	P62273	Structural constituent of ribosome Cellular nitrogen compound and protein metabolic processes
Proteins declining in hepatocyte-like cells (Day 40)		
Glycine amidinotransferase, mitochondrial	P50440	Creatinine biosynthesis Cellular nitrogen compound and protein metabolic processes

Protein description	Main accession	Molecular function or biological process involvement
Proteins declining in hepatocyte-like cells (Day 40)		
ADP-ribosylation factor-like protein 8B	Q9NVJ2	Lysosome motility Tubulin binding Cell cycle and chromosome segregation
Peroxisomal membrane protein 11B	O96011	Peroxisome proliferation and division
Fumarylacetoacetase	P16930	Catabolic processes L-phenylalanine degradation (amino acid degradation)
Apolipoprotein A-IV	P06727	Activation of lipoprotein lipase Cholesterol binding and transport

Hierarchical clustering was done using averaged group data (Figure 4.15). Generating ten row clusters for individual samples produced clusters ranging from 1 to 5 290 proteins per cluster. Intermediately sized clusters included 38, 46, 71 and 126 proteins. Small groups demonstrated no trends while clusters containing 43 and 71 proteins decreased at Day 36 or from Day 16 (Figure 4.16) respectively.

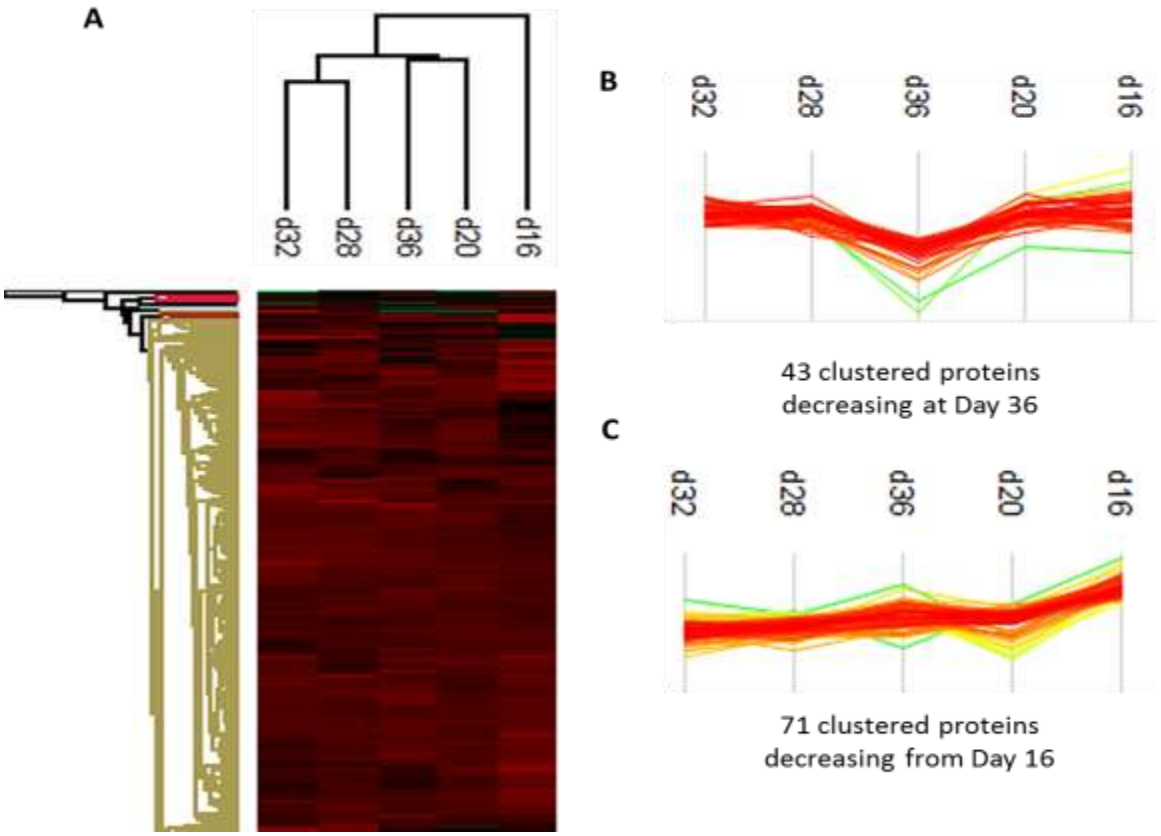

Figure 4.15: A) Hierarchical clustering through the HLC maturation time course from Day 16 to Day 36 and B - C) Protein profiles of clusters demonstrating interesting trends (proteins listed in appendix).

Samples from Day 16 to Day 36 did not co-segregate according to the stage of differentiation. The use of volcano plots for easy identification of changes in large comparative datasets is commonly used in proteomic analysis. Volcano plots show statistical significance (-Log p-value) versus mean difference in large replicate datasets. This enables visualisation of differences plotted as changes in both directions equidistant from the centre. Low p-values, therefore highly significant values, appear to the top of plots. The relative importance of the t-test value and the difference between the mean can be adjusted in relation to the data set size.

Volcano plots were generated using a t-test with 250 randomizations at an FDR of 0.01 using a weighting for differences in the mean (s0 = 0.1) (Tusher VG, Tibshirani R *et al.*, 2001). The variance between samples using volcano plots (Figure 4.16) highlighted the similarity of the global proteome of these samples throughout the maturation phase of differentiation. Detailing the maturation of the hepatic proteome using hierarchical clustering and volcano plots provided little insight into the global proteomic changes during maturation. This would necessitate further statistical rigor in assessing specific hepatic proteins of interest extracted from the data set.

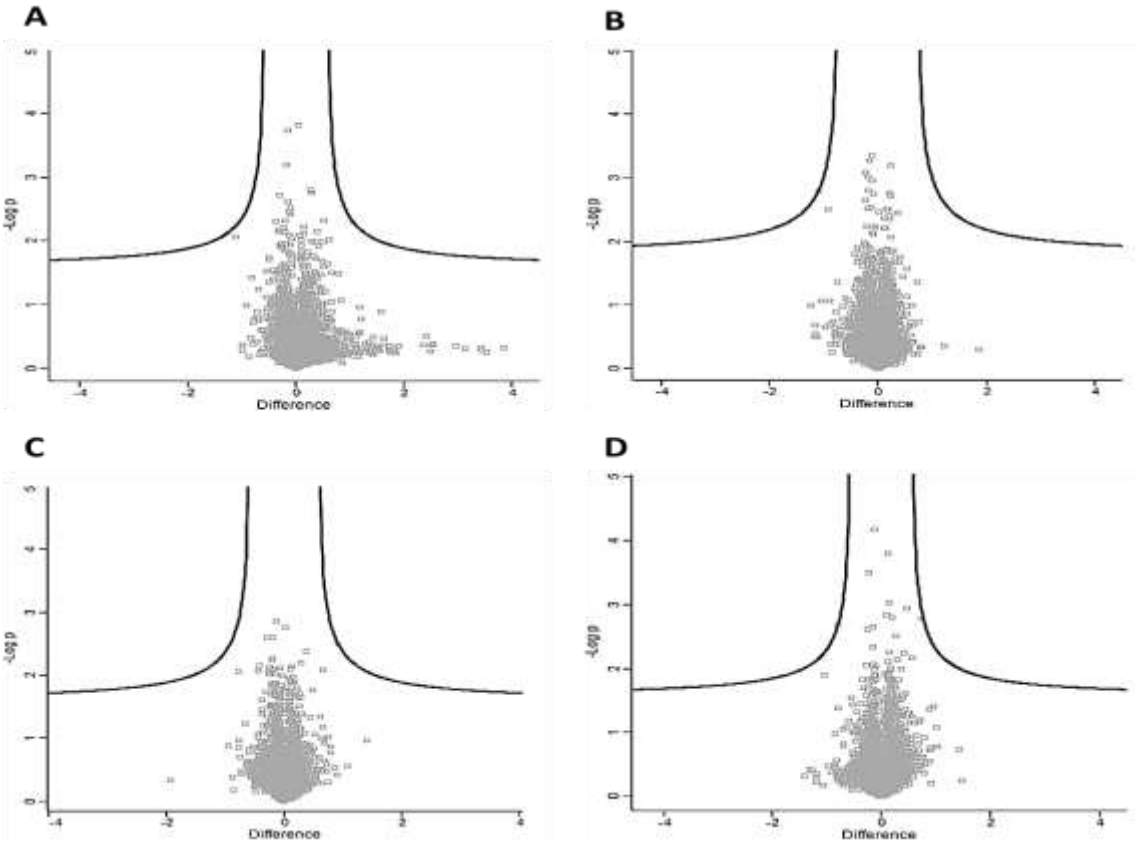

Figure 4.16: Volcano plots comparing different time points using 250 randomizations, an FDR of 0.05 with contributions from the p-value and the mean difference. A) Day 16 versus Day 20, B) Day 20 versus Day 28, C) Day 28 versus Day 32 and D) Day 32 versus Day 36.

Hepatic differentiation of mesenchymal stem cells (MSCs) in a two-step differentiation protocol (HGF and FGF for 7 days followed by oncostatin M, dexamethasone, insulin, transferrin and selenium) demonstrated hepatocyte maturity after 6 weeks which was retained over 12 weeks (Lee KD, Kuo TKC *et al.*, 2004). Since then, much research has been done to optimise differentiation into HLCs. The proteome generated in this study, based on the protocols of Gieseck *et al.* (2015), suggests a window of hepatic phenotype maturity which declines within approximately 5 weeks from induction. Markers more exclusive to a foetal phenotype such as CYP3A7 were not detected. Due to the overlap in proteins which define foetal and mature adult phenotypes it was difficult to draw a definitive correlation to foetal or adult hepatic maturity.

4.3 Conclusion

Others have shown that various markers, both at the transcript and protein level, are comparable between HLCs and freshly isolated PHHs. While other proteins are poorly mimicked by HLCs. HLCs may be limited in differentiation homogeneity, resulting in variable maturation and expression of the repertoire of metabolising enzymes essential to hepatocyte function. However, if HLC maturity can be established to fall within well-defined constraints, then iPSC-derived hepatocytes may replace PHHs as a reproducible and feasible hepatic model system. In addition, benchmarking to mature hepatocytes is required to demonstrate the validity and applicability of these cell for use in safety pharmacology.

The data presented here suggests that the value in differentiation of iPSCs into HLCs with maintenance of hepatic markers could be between Day 30 and Day 32. The strength of this specific differentiation protocol can be seen by the lengths to which HLCs have progressed from undifferentiated precursors. In this data set, limited replicates could result in over-interpretation of the systems represented. Pooling of multiple samples from the same time point prior to labelling may have added to the biological diversity of this analysis and provided more definitive data.

Despite the limited replicates, proteins corresponding to the functional repertoire of the liver were expressed in a time-dependent manner and peaked at Day 30 and Day 32 respectively in each time course. This is a positive indication of the acquisition of a hepatic phenotype but fold-changes in relative ratios of hepatic marker proteins did not pass basic fold-change cut-offs during maturation. Unfortunately, this data does not meet the requirements for more rigorous statistical analysis and therefore definitive confirmation of an adequate end-point for differentiation is not possible. The reproducibility of differentiation and degree of maturation may better have been assessed with fewer time-points with more replicates around day 30 to 36 as opposed to spanning 20 days of maturation.

4.4 Appendix: Characterisation of Proteome Changes during Hepatocyte Differentiation

Table 4.16: Proteins from HLCTC hierarchical clustering demonstrating interesting trends

Cluster 1 (proteins increasing until Day 35)		Cluster 2 (proteins increasing until Day 30)	
Accession	**Description**	**Accession**	**Description**
Q9BPX1	17-beta-hydroxysteroid dehydrogenase 14	Q7L266	Isoaspartyl peptidase/L-asparaginase
Q8TF42	Ubiquitin-associated and SH3 domain-containing protein B	P23677	Inositol-trisphosphate 3-kinase A
Q7Z7D3	V-set domain-containing T-cell activation inhibitor 1	P01031	Complement C5
P52566	Rho GDP-dissociation inhibitor 2	P00995	Serine protease inhibitor Kazal-type 1
P27487	Dipeptidyl peptidase 4	O94768	Serine/threonine-protein kinase 17B
P17301	Integrin alpha-2	O43772	Mitochondrial carnitine/acylcarnitine carrier protein
P17936	Insulin-like growth factor-binding protein 3	O00115	Deoxyribonuclease-2-alpha
Q96HE7	ERO1-like protein alpha	O14967	Calmegin
Q8IWE2	Protein NOXP20	P04066	Tissue alpha-L-fucosidase
P49441	Inositol polyphosphate 1-phosphatase	Q68DH5	LMBR1 domain-containing protein 2
P30740	Leukocyte elastase inhibitor	Q9UJM3	ERBB receptor feedback inhibitor 1
P21291	Cysteine and glycine-rich protein 1	Q9UBR2	Cathepsin Z
P13473	Lysosome-associated membrane glycoprotein 2	Q96FQ6	Protein S100-A16
P05121	Plasminogen activator inhibitor 1	Q5T2W1	Na(+)/H(+) exchange regulatory cofactor NHE-RF3
P04179	Superoxide dismutase [Mn], mitochondrial	P21266	Glutathione S-transferase Mu 3
P04083	Annexin A1	P12429	Annexin A3
O95070	Protein YIF1A	P02751	Fibronectin
O95834	Echinoderm microtubule-associated protein-like 2	P09327	Villin-1
Q9Y6N5	Sulfide:quinone oxidoreductase, mitochondrial	Q9UHL4	Dipeptidyl peptidase 2
Q9Y646	Carboxypeptidase Q	Q9NP58	ATP-binding cassette sub-family B member 6
Q9HAT2	Sialate O-acetylesterase	Q9HCY8	Protein S100-A14
Q96NY8	Nectin-4	Q99956	Dual specificity protein phosphatase 9
Q96I99	Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial	P42330	Aldo-keto reductase family 1 member C3
Q8WTV0	Scavenger receptor class B member 1	P30039	Phenazine biosynthesis-like domain-containing protein
Q8IV56	Proline-rich protein 15	P25815	Protein S100-P
Q15582	Transforming growth factor-beta-induced protein	P09110	3-ketoacyl-CoA thiolase, peroxisomal
Q13443	Disintegrin and metalloproteinase domain-containing protein 9	P07858	Cathepsin B
Q12864	Cadherin-17	P06727	Apolipoprotein A-IV
P40261	Nicotinamide N-methyltransferase	P02771	Alpha-fetoprotein
P22309	UDP-glucuronosyltransferase 1-1	P02766	Transthyretin
P21695	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	P02679	Fibrinogen gamma chain
O43581	Synaptotagmin-7	P02671	Fibrinogen alpha chain
P18440	Arylamine N-acetyltransferase 1	P02647	Apolipoprotein A-I
Q9H9S4	Calcium-binding protein 39-like	P01019	Angiotensinogen
Q9H8M7	Protein FAM188A	P01009	Alpha-1-antitrypsin
Q9BUP0	EF-hand domain-containing protein D1	P00738	Haptoglobin
Q96C19	EF-hand domain-containing protein D2	O75503	Ceroid-lipofuscinosis neuronal protein 5

Cluster 1 (proteins increasing until Day 35)		Cluster 2 (proteins increasing until Day 30)	
Q7L5N7	Lysophosphatidylcholine acyltransferase 2	O95210	Starch-binding domain-containing protein 1
Q53EL6	Programmed cell death protein 4	Q9H008	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase
Q07352	Zinc finger protein 36, C3H1 type-like 1	Q9BZV2	Thiamine transporter 2
Q02809	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	Q9BSF4	Uncharacterized protein C19orf52
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	Q8NBW4	Sodium-coupled neutral amino acid transporter 9
P51608	Methyl-CpG-binding protein 2	Q06520	Bile salt sulfotransferase
P00813	Adenosine deaminase	P02760	Protein AMBP
P22676	Calretinin	P00167	Cytochrome b5
Q9Y624	Junctional adhesion molecule A	P02656	Apolipoprotein C-III
Q96TA1	Niban-like protein 1	P31327	Carbamoyl-phosphate synthase [ammonia], mitochondrial
Q8NHG8	E3 ubiquitin-protein ligase ZNRF2	P22732	Solute carrier family 2, facilitated glucose transporter member 5
Q8IVL1	Neuron navigator 2	P21549	Serine--pyruvate aminotransferase
Q86TA1	MOB kinase activator 3B	P15144	Aminopeptidase N
Q4KWH8	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase-1	P07148	Fatty acid-binding protein, liver
Q14894	Ketimine reductase mu-crystallin	P07306	Asialoglycoprotein receptor 1
Q12986	Transcriptional repressor NF-X1	Q9Y2S2	Lambda-crystallin homolog
Q08380	Galectin-3-binding protein	Q9ULC5	Long-chain-fatty-acid--CoA ligase 5
Q01415	N-acetylgalactosamine kinase	P55157	Microsomal triglyceride transfer protein large subunit
P54802	Alpha-N-acetylglucosaminidase	P21980	Protein-glutamine gamma-glutamyltransferase 2
P37235	Hippocalcin-like protein 1	P21964	Catechol O-methyltransferase
P26447	Protein S100-A4	P18827	Syndecan-1
O95571	Persulfide dioxygenase ETHE1, mitochondrial	P05062	Fructose-bisphosphate aldolase B
P08962	CD63 antigen	P17931	Galectin-3
Q9BZ29	Dedicator of cytokinesis protein 9	Q9H6S3	EGF receptor kinase substrate 8-like protein 2
Q07075	Glutamyl aminopeptidase	Q96JY6	PDZ and LIM domain protein 2
P18564	Integrin beta-6	Q92597	Protein NDRG1
P16930	Fumarylacetoacetase	Q15493	Regucalcin
P08236	Beta-glucuronidase	Q02318	Sterol 26-hydroxylase, mitochondrial
P02794	Ferritin heavy chain	P00352	Retinal dehydrogenase 1
P07996	Thrombospondin-1	P50120	Retinol-binding protein 2
		Q8TD22	Sideroflexin-5

Table 4.17: Proteins from HLCLTC hierarchical clustering demonstrating interesting trends

Cluster 1 (proteins decreasing from Day 16)		Cluster 2 (proteins decreasing at Day 36)	
Accession	**Description**	**Accession**	**Description**
Q9UKM9	RNA-binding protein Raly	Q9H6R4	Nucleolar protein 6
Q9H9Y6	DNA-directed RNA polymerase I subunit RPA2	Q96HH9	GRAM domain-containing protein 3
Q96QD9	UAP56-interacting factor	Q92828	Coronin-2A
Q96E39	RNA binding motif protein, X-linked-like-1	Q8IY81	pre-rRNA processing protein FTSJ3
Q92959	Solute carrier organic anion transporter family member 2A1	Q6P6B1	Glutamate-rich protein 5
Q5BKZ1	DBIRD complex subunit ZNF326	Q15651	High mobility group nucleosome-binding domain-containing protein 3
P33991	DNA replication licensing factor MCM4	Q08380	Galectin-3-binding protein
P02654	Apolipoprotein C-I	O00479	High mobility group nucleosome-binding domain-containing protein 4
P04792	Heat shock protein beta-1	P14384	Carboxypeptidase M
Q9Y294	Histone chaperone ASF1A	Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog
Q9NZU0	Leucine-rich repeat transmembrane protein FLRT3	Q9Y2X3	Nucleolar protein 58
Q9NTJ3	Structural maintenance of chromosomes protein 4	Q9NY12	H/ACA ribonucleoprotein complex subunit 1
Q9NRZ5	1-acyl-sn-glycerol-3-phosphate acyltransferase delta	P68431	Histone H3.1
Q9BT25	HAUS augmin-like complex subunit 8	P62805	Histone H4
Q8WZA9	Immunity-related GTPase family Q protein	P55769	NHP2-like protein 1
Q8WV93	Lactation elevated protein 1	P52926	High mobility group protein HMGI-C
Q8NBI6	Xyloside xylosyltransferase 1	P22087	rRNA 2'-O-methyltransferase fibrillarin
Q5VZL5	Zinc finger MYM-type protein 4	O76021	Ribosomal L1 domain-containing protein 1
Q5SZK8	FRAS1-related extracellular matrix protein 2	O75367	Core histone macro-H2A.1
Q3SY77	UDP-glucuronosyltransferase 3A2	O60832	H/ACA ribonucleoprotein complex subunit 4
Q12923	Tyrosine-protein phosphatase non-receptor type 13	O00124	UBX domain-containing protein 8
Q01974	Tyrosine-protein kinase transmembrane receptor ROR2	O00567	Nucleolar protein 56
Q00534	Cyclin-dependent kinase 6	Q9Y5L0	Transportin-3
P62166	Neuronal calcium sensor 1	Q9UK45	U6 snRNA-associated Sm-like protein LSm7
P42702	Leukemia inhibitory factor receptor	Q9NX24	H/ACA ribonucleoprotein complex subunit 2
P23508	Colorectal mutant cancer protein	Q9BVJ6	U3 small nucleolar RNA-associated protein 14 A
P13726	Tissue factor	Q9BR61	Acyl-CoA-binding domain-containing protein 6
O95347	Structural maintenance of chromosomes protein 2	Q6FIF0	AN1-type zinc finger protein 6
P02462	Collagen alpha-1(IV) chain	Q2TBE0	CWF19-like protein 2
Q8TDP1	Ribonuclease H2 subunit C	Q15427	Splicing factor 3B subunit 4
Q5KU26	Collectin-12	P52815	39S ribosomal protein L12, mitochondrial
Q14781	Chromobox protein homolog 2	P22415	Upstream stimulatory factor 1
Q14566	DNA replication licensing factor MCM6	O43715	TP53-regulated inhibitor of apoptosis 1
Q13257	Mitotic spindle assembly checkpoint protein MAD2A	P12074	Cytochrome c oxidase subunit 6A1, mitochondrial

Cluster 1 (proteins decreasing from Day 16)		Cluster 2 (proteins decreasing at Day 36)	
P49736	DNA replication licensing factor MCM2	Q9H190	Syntenin-2
P33993	DNA replication licensing factor MCM7	Q9H0W9	Ester hydrolase C11orf54
P33992	DNA replication licensing factor MCM5	Q96MV1	Transmembrane protein 56
P29373	Cellular retinoic acid-binding protein 2	Q14749	Glycine N-methyltransferase
P25205	DNA replication licensing factor MCM3	O76054	SEC14-like protein 2
P18858	DNA ligase 1	P02774	Vitamin D-binding protein
P13611	Versican core protein	Q9ULC5	Long-chain-fatty-acid--CoA ligase 5
O75717	WD repeat and HMG-box DNA-binding 1	Q8N5J2	Protein FAM63A
P09486	SPARC	P62266	40S ribosomal protein S23
Q9UPY5	Cystine/glutamate transporter	P30039	Phenazine biosynthesis-like domain-containing protein
Q9NWM0	Spermine oxidase (	P16104	Histone H2AX
Q96N64	PWWP domain-containing protein 2A	P21964	Catechol O-methyltransferase
Q6PCE3	Glucose 1,6-bisphosphate synthase		
Q5TFE4	5'-nucleotidase domain-containing protein 1		
Q2Q1W2	E3 ubiquitin-protein ligase TRIM71		
Q15024	Exosome complex component RRP42		
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1		
Q02539	Histone H1.1		
P48681	Nestin		
P31483	Nucleolysin TIA-1 isoform p40		
P11169	Solute carrier family 2, facilitated glucose transporter member 3		
P08670	Vimentin		
O14495	Phospholipid phosphatase 3		
O75494	Serine/arginine-rich splicing factor 10		
Q8NFJ5	Retinoic acid-induced protein 3		
Q0VAQ4	Small cell adhesion glycoprotein		
P02511	Alpha-crystallin B chain		
P48539	Purkinje cell protein 4		
Q9UHI5	Large neutral amino acids transporter small subunit 2		
Q9NZU5	LIM and cysteine-rich domains protein 1		
Q96D15	Reticulocalbin-3		
Q4L180	Filamin A-interacting protein 1-like		
Q14451	Growth factor receptor-bound protein 7		
Q01740	Dimethylaniline monooxygenase [N-oxide-forming] 1		
P63302	Selenoprotein W		
O95340	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase 2		
P40261	Nicotinamide N-methyltransferase		

Chapter 5: Proteomic Comparison of Hepatocytes

5.1 Materials and Methods

5.1.1 Gel-based Proteomics

Materials and methods for sample collection, protein preparation and one-dimensional SDS-PAGE were conducted as described in Chapter 4.

5.1.2 Mass Spectrometry-based Proteomics

Materials and methods for preparing proteins from cultured cells including; sample collection and lysis, protein quantitation, protein digestion, peptide TMT labelling, solid phase extraction clean-up, bulk reverse phase UPLC with fractionation and selected recombination and the final LC-MS/MS analysis were conducted in an identical manner to that as described in Chapter 4. Replicates of each sample or growth condition were labelled with different TMTs to avoid any experimental or reaction bias and to give some idea of the technical variance of the methodology (Table 5.1). TMT-labelled peptides for the comparison of hepatocyte proteomes were run at the Cambridge Centre for Proteomics Core Facility (Department of Biochemistry, University of Cambridge).

Table 5.1: Tandem mass tags used to label the proteome of hepatocytes from various origins and different culturing techniques

TMT Label	Replicate 1	Replicate 2	Replicate 3
126	HepG2_3D (Day 10)	HLC (Day 35)	HLC (Day 35)
127	PHH	HLC (Day 35)	HLC (Day 35)
128	HepG2 monolayer	HepG2 monolayer	HepG2 monolayer
129	HepG2 monolayer	PHH	PHH
130	HLC (Day 35)	PHH	HepG2_3D (Day 10)
131	HLC (Day 35)	HepG2_3D (Day 10)	HepG2_3D (Day 14)

5.2 Results and Discussion

5.2.1 Proteomic Comparison of Hepatocytes

Benchmarking HLC maturity can be done in comparison to various hepatic tissue, the choice of which dictates the perceived degree of maturation. Known gene markers of hepatic maturity and metabolic capacity (ALB, APOF, HNF4α, CYP1A2, CYP3A4, CYP3A5 and CYP3A7) have been shown to be reduced in plated cells compared to fresh PHHs frozen upon isolation. In other studies, freshly isolated PHHs were considered the best benchmark for gene expression of functional hepatocytes (Segeritz C, 2015) and provided support for the use of cryopreserved, pooled PHH for proteomic comparisons.

Prior to MS-based proteomics, protein mass profiles of each of the replicate samples, were generated using SDS-PAGE to highlight any general band differences between the different hepatocyte lysates and to ensure lysis efficiency. The integrity of the different samples was confirmed and the differences throughout the mass profiles between PHHs and HepG2 cells (Figure 5.1) as well as between the HepG2 spheroids, HepG2 monolayers and PHHs (Figure 5.2) noted.

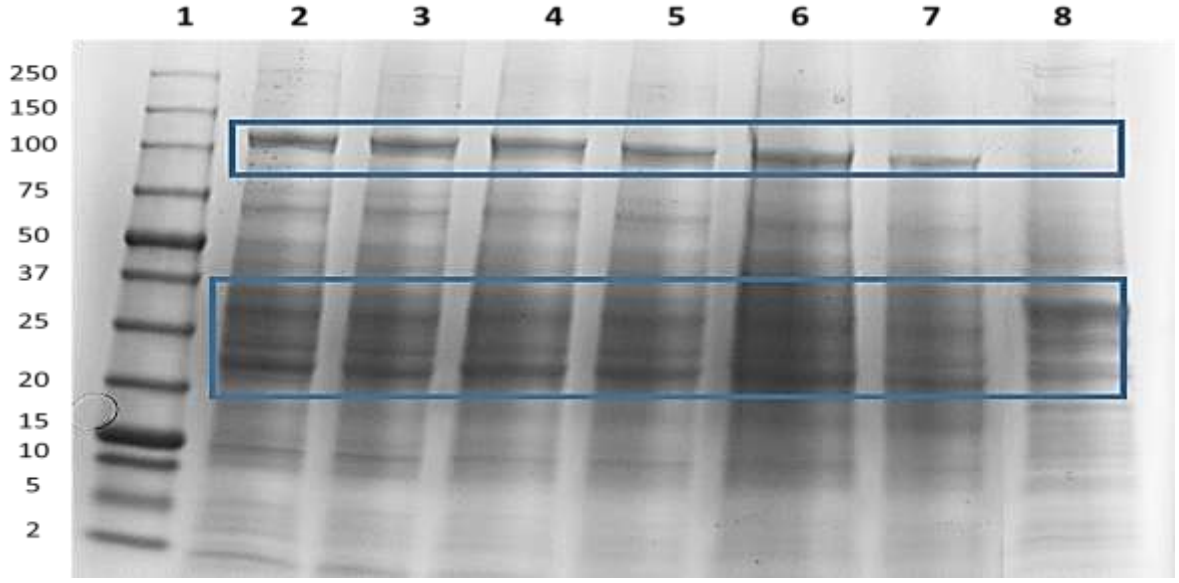

Figure 5.1: Coomassie stained 4-15% Mini-PROTEAN TGX gel of hepatocyte lysates. 1) Precision Plus Protein Dual Colour standard (2 - 250 kDa), 2 - 7) Pooled primary human hepatocytes collected four hours post-thawing from six different wells and 8) HepG2 cells collected at ten days post-thawing. Blue boxes highlight mass regions with distinct differences.

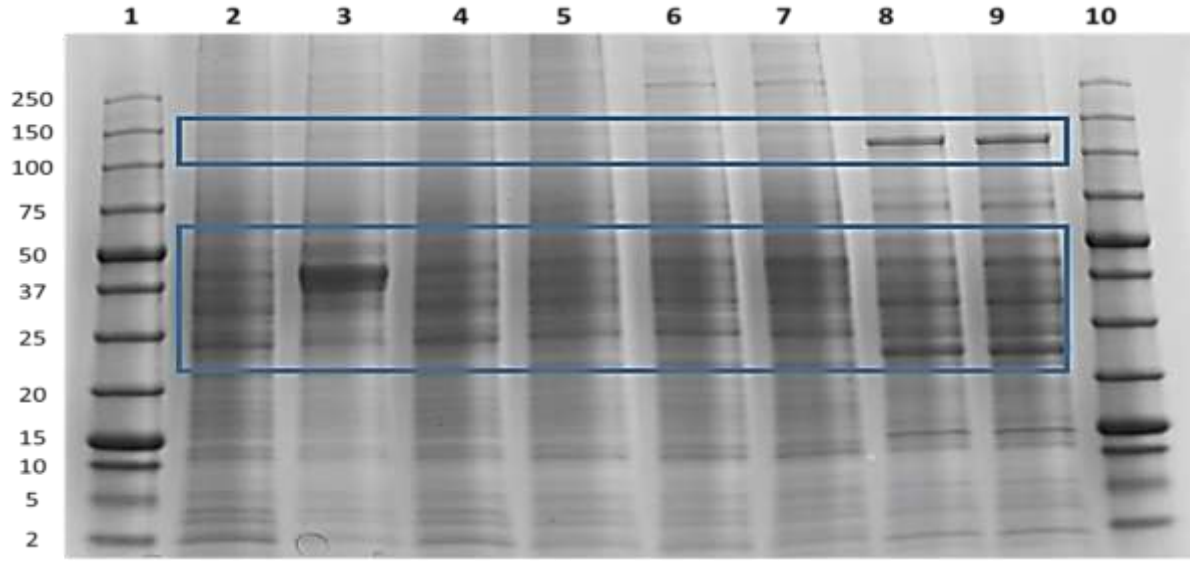

Figure 5.2: Coomassie stained 4-15% Mini-PROTEAN TGX gel of hepatocyte lysates. 1) Precision Plus Protein Dual Colour standard (2 - 250 kDa), 2) HepG2 cell spheroids collected at Day 10, 3) HepG2 cell spheroids collected at Day 14, 4) HepG2 cell spheroids collected at Day 10, 5 - 7) HepG2 cell monolayers at various passages, 8 - 9) Primary human hepatocytes and 10) Standard (2 – 250 kDa). Blue boxes highlight mass regions with distinct differences.

After confirmation of hepatocyte lysis and successful protein profiling, the samples were prepared for labelling and mass spectrometric analysis following which the raw data files were processed as described in Chapter 4. The number of proteins identified for the hepatocyte comparisons (HLCA) was 6 682, 6 285 and 6 449 for replicates 1 to 3 respectively. Results were again filtered to include only proteins with 2 or more unique peptides at 100% identification confidence which produced 5 231 proteins for further interrogation (Table 5.2).

Table 5.2: Global summary for replicates of hepatocyte comparisons

	Hepatocyte comparison (HLCA)		
	Replicate 1	**Replicate 2**	**Replicate 3**
Proteins	6 682	6 285	6 449
Protein group(s): 2+ unique peptides	6 511	6 099	6 261
Protein group(s): 2+ unique peptides, high confidence (100%)	6 387	5 936	6155
	Protein group(s): 2+ unique peptides, high confidence		
Range of peptide numbers	2 - 371	2 - 483	2 -374
Range of peptide spectrum matches	2 - 966	2 - 1504	2 - 896
Molecular weight range (kDa)	5.04 - 1010.45	5.02 - 860.12	5.04 - 3813.65
Coverage range	0.17 - 88.24	0.16 - 91.67	0.13 - 91.67
Proteins in replicate 1 to 3	5 231		
Proteins in replicate 1 and 2	277		
Proteins in replicate 1 and 3	472		
Proteins in replicate 2 and 3	180		
Proteins in replicate 1 only	407	-	-
Protein in replicate 2 only	-	248	-
Protein in replicate 3 only	-	-	272

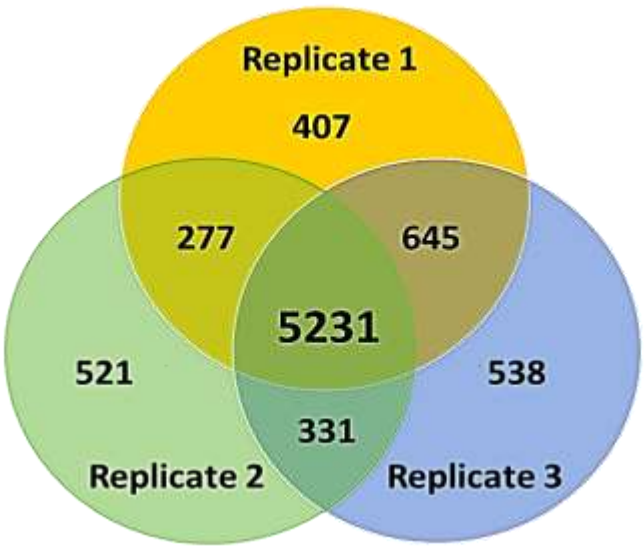

Figure 5.3 Venn diagram illustrating the overlap of proteins identified across replicates.

Histograms and summary statistics of coverage, molecular weight distribution, number of peptides, PSMs and isobaric tags exhibit similarity across equivalent replicates (Figures 5.4 to 5.7). The notable similarity combined with the high overlap in the number of proteins identified is suggestive of high quality acquisition and scientific data processing.

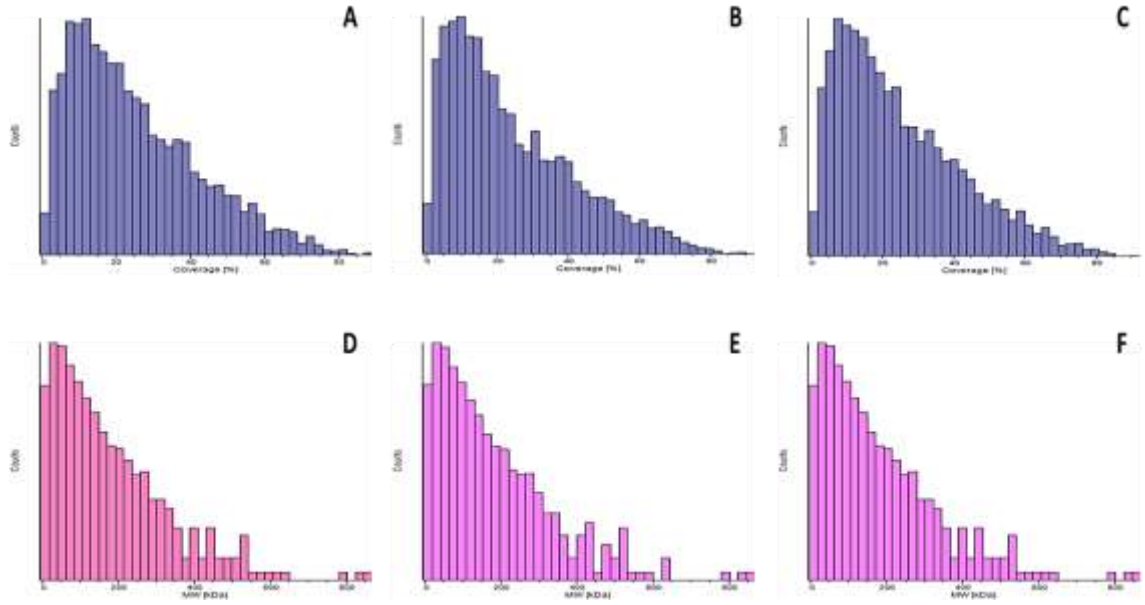

Figure 5.4: Histograms illustrating, A - C) Protein coverage (%) in replicates 1 to 3 and D - F) Molecular weight (kDa) range of proteins in replicates 1 to 3 (x-axis: counts).

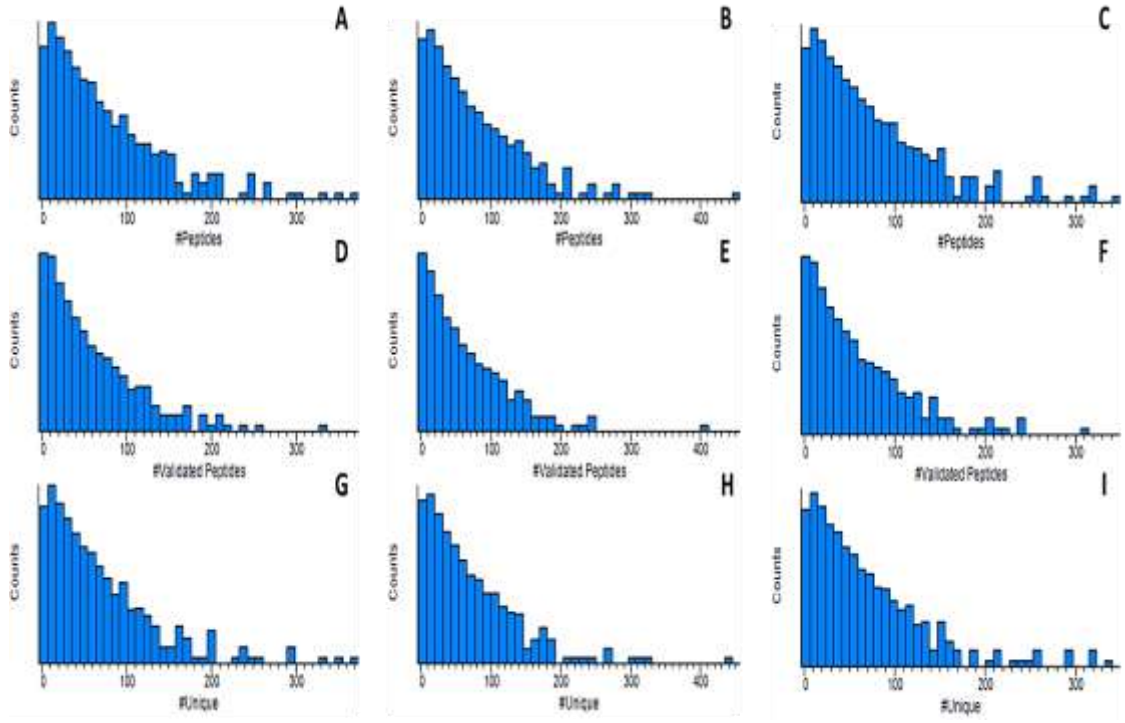

Figure 5.5: Histograms of HLCA replicates 1 to 3 for A - C) Number of peptides, D - F) Number of validated peptides and G - I) Number of unique peptides (x-axis: counts).

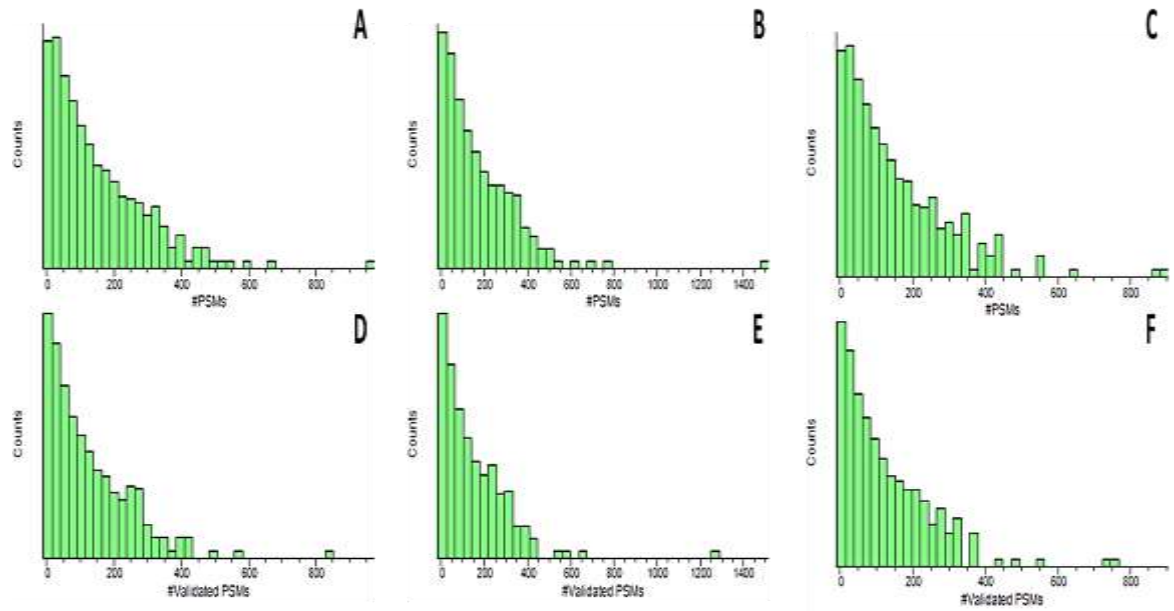

Figure 5.6: Histograms of HLCA replicates 1 to 3 for A - C) Number of peptide-spectrum matches and D - F) Number of validated peptide-spectrum matches (x-axis: counts).

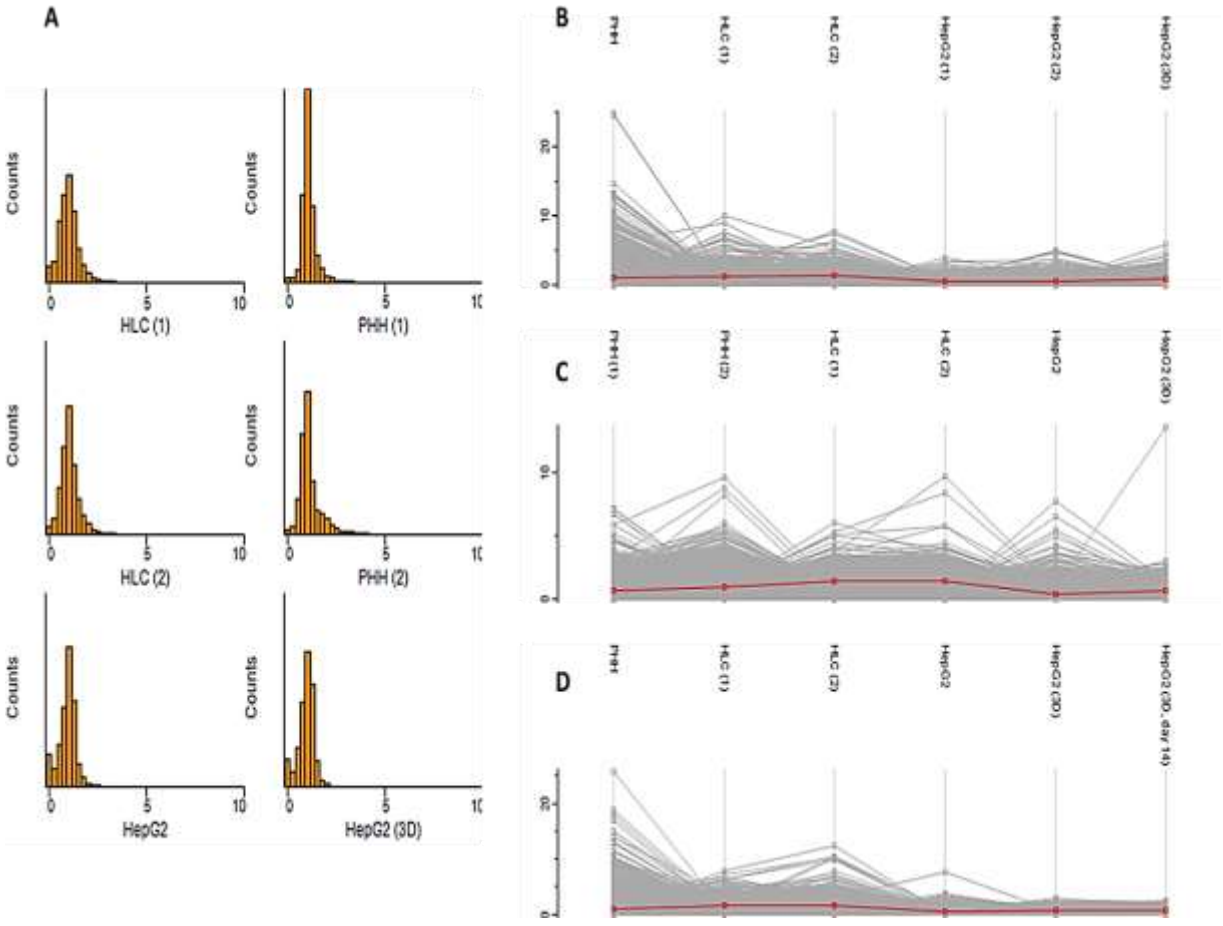

Figure 5.7: Tandem mass tag distribution and protein profiles. A) Histogram of HLCA replicate 1 illustrating TMT distribution, B - D) Average protein profiles for replicate 1 to 3 respectively. Red line on protein profiles: glyceraldehyde-3-phosphate dehydrogenase (GAPDH; P04406).

5.2.2 General Overview of Comparisons

Whole genome expression profiling of PHHs, HepG2 and Huh7 cells, done by others, demonstrates poor correlation with human liver tissue. However, PHHs in sandwich culture with rBM possess liver-specific markers and *in vivo* functional responses. Lack of correlation can be partially attributed to the presence of Kupffer and hepatic stellate cells in whole liver tissue. Notably, increased AFP, indicative of a loss of mature liver phenotype, is found in HepG2 and Huh7 cells. Furthermore, expression of some CYP450 isoforms, phase II and III transporters and overall biotransformation responses are absent. This demonstrates the limited potential of hepatoma derived cell lines in monolayer cultures at simulating the *in vivo* hepatic phenotype (Olsavsky KM, Page JL *et al.*, 2007).

Correlating whole genome expression to the whole proteome could aid in further defining the advantages and limitations of these hepatic models. Using quantitative proteomic data together with protein identity allows several bioanalytical assays to be performed and data to be interrogated for specific aspects of global biological processes that take place within the different tissue. One technique that can quickly highlight differences in groups of compared samples is hierarchical clustering.

Hierarchical clustering was used to assess the different hepatic cells and culturing conditions used in this study (Figure 8A) which appeared to distinctively grouped PHHs, HLCs and HepG2 cells. Hierarchical clustering was also done using the average of biological replicates (Figure 5.8C). Generating ten row clusters for individual samples produces clusters ranging from 2 to 3 682 proteins per cluster.

Samples generally co-segregated according to cell type with HepG2 cells in monolayers and spheroids segregating together. HLCs and PHHs co-segregated separately with one exception being PHH3 which clustered with HepG2 spheroids. This could be due to various factors including biological variance associated with thawing PHH, insufficient quantity of protein labelled or reduced labelling efficiency. As the reason for this clustering could not definitively established, this data could not be excluded.

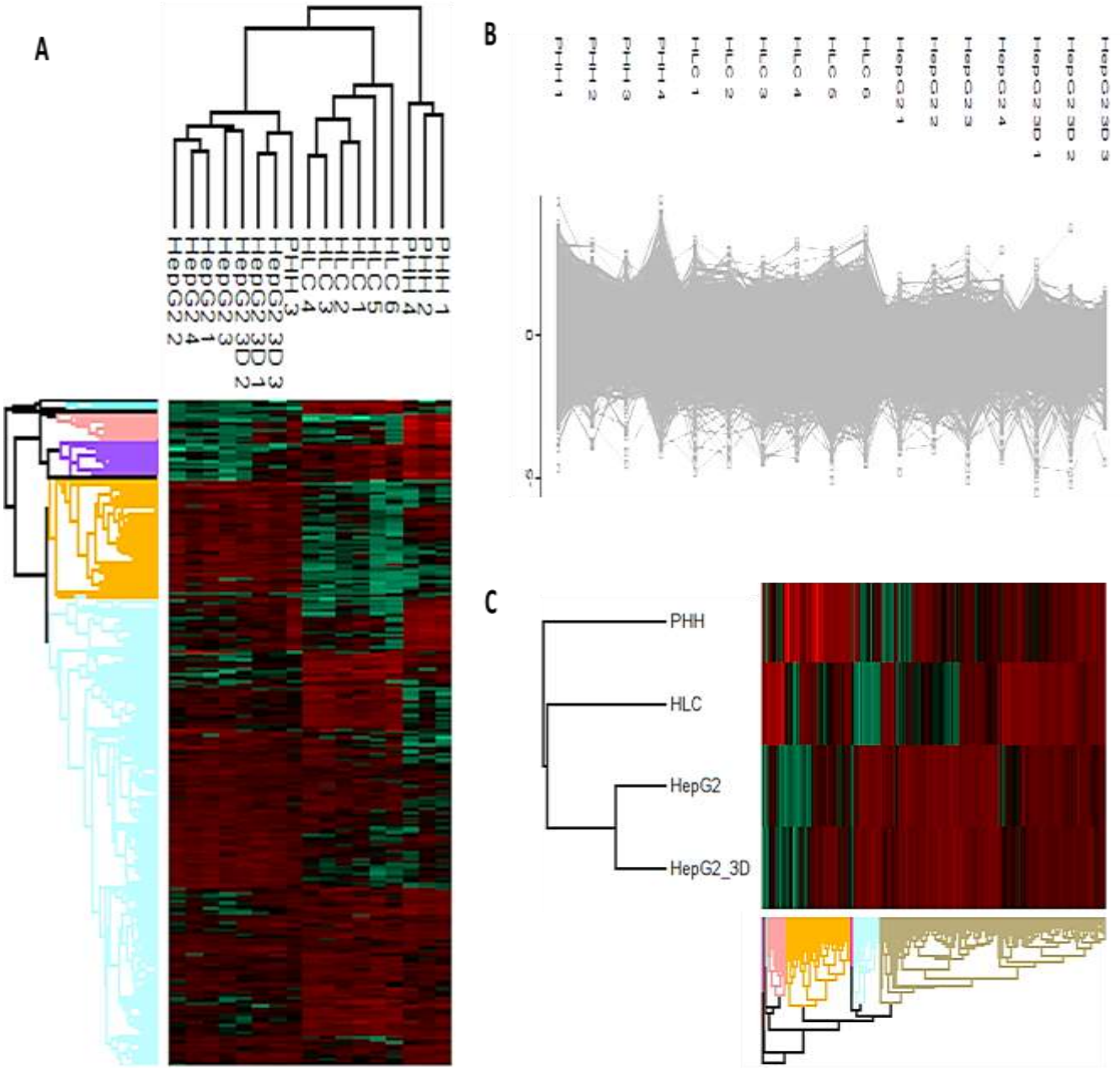

Figure 5.8: A) Hierarchical clustering of proteome data for individual samples defined with ten row clusters, B) Total protein cluster profile and C) Hierarchal clustering of averaged replicate protein abundance data according to cell type or culturing technique.

Hierarchical clustering demonstrated distinct patterning with clustering that could be established and interrogated as required (Figure 5.9). Clustering demonstrated circumstances under which protein expression was greater in HLCs and PHHs compared to HepG2 cells (Cluster 4). And also where HLCs were distinguished from all the other cell types (Cluster 2, 4 and 6). This broad row clustering enabled an overview of the similarity of cell types but did not provide any detail as the type of proteins displaying similarities or differences.

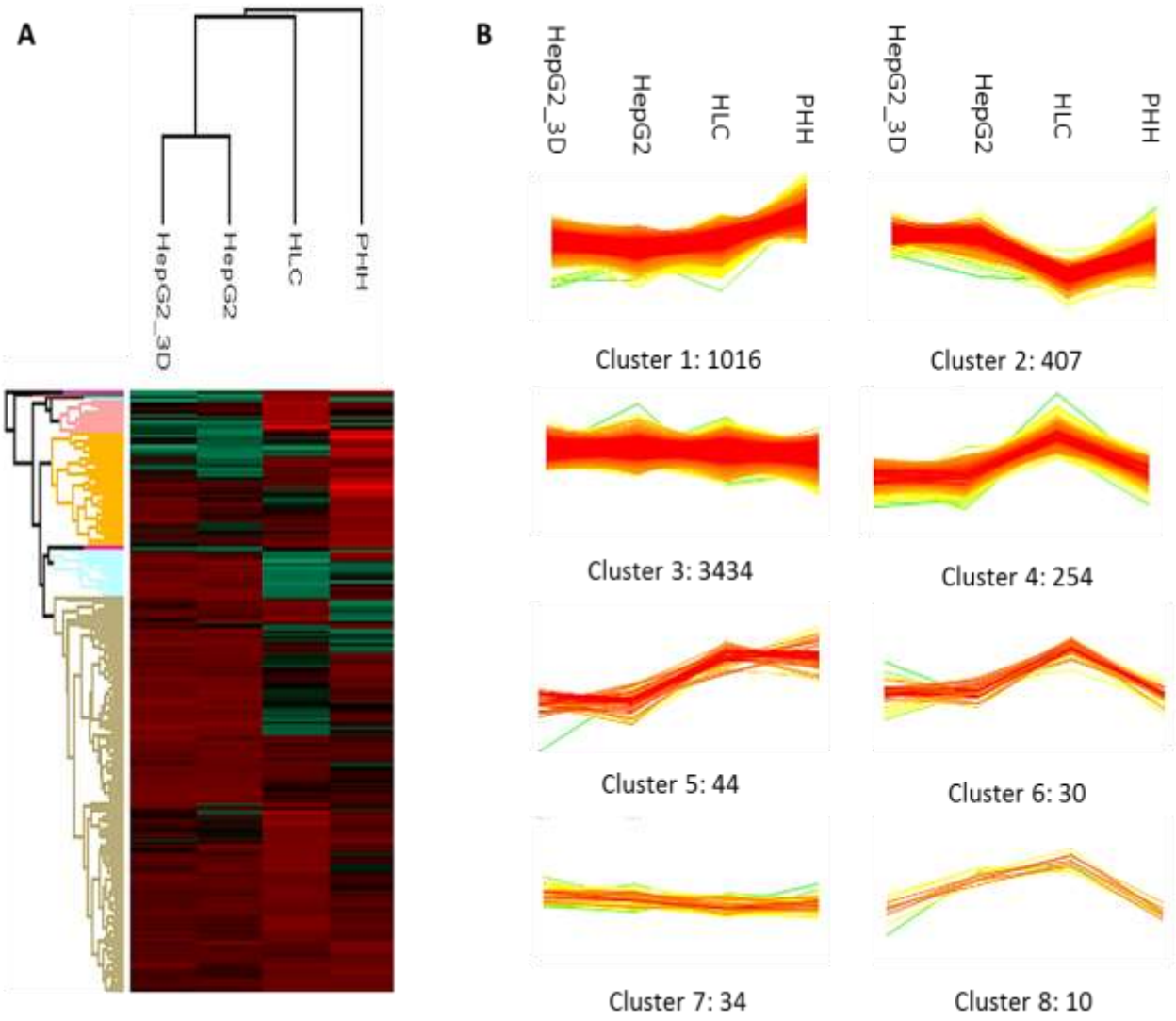

Figure 5.9: A) Hierarchical clustering of HLCA biological replicates and B) Protein profiles produced from 10 row clusters with the corresponding number of associated proteins. Clusters 9 and 10 contained a single protein only.

Principal component analysis (PCA), clustered PHHs (with the exception of PHH3) and HLCs distinctly in Component 1 (40.5%) and Component 2 (33.2%) of principal component space (Figure 5.10A). In contrast, HepG2 cells, regardless of culturing strategy, clustered together. When looking at the lower principal component which account for less of the overall variation, Component 3 (4.8%), Component 4 (4.3%) and Component 5 (2.7%) spatially resolved HepG2 monolayers and spheroids (Figure 5.10B). Despite not being very tightly clustered, trends in data were still apparent.

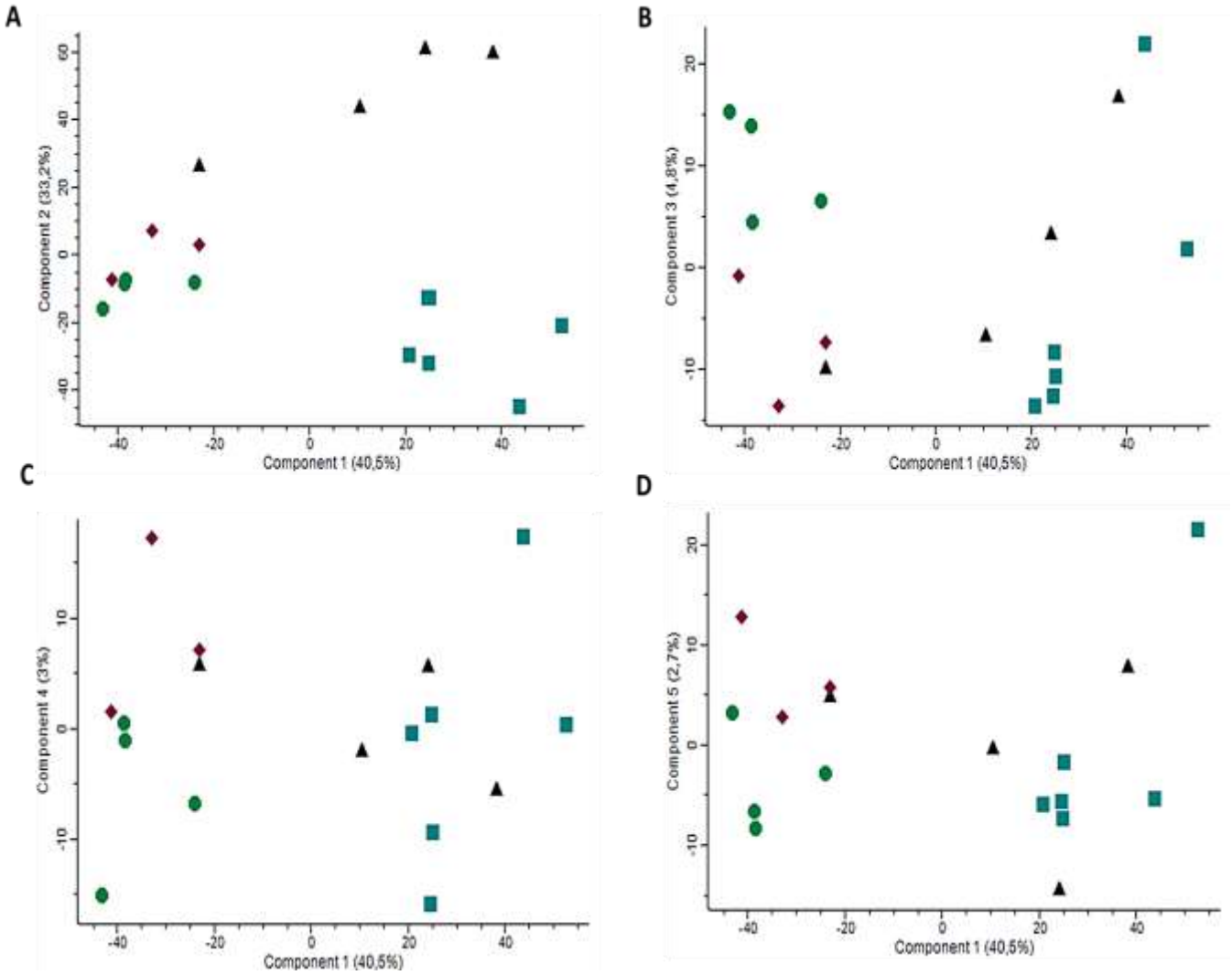

Figure 5.10: Principal component analysis (PCA) scatter plots of A) Component 1 versus Component 2, B) Component 1 versus Component 3, C) Component 1 versus Component 4 and D) Component 1 versus Component 5. PHHs: black triangle (n = 4), HLCs: blue square (n = 6), HepG2 monolayers: green dot (n = 4), HepG2 spheroids: purple diamond (n = 3).

In the lower principal components, PHHs and HLCs clustered similarly and often interspersed with one another. HepG2 monolayers and spheroids, along with PHH3, were found behave similarly and remain separate from PHHs and HLCs. These trends continued from Component 6 (2.6%) through to Component 16 (0.6%). The principle component clustering as illustrated in Figure 5.8 shows relatively weak but exclusive clustering depending on which components were being compared.

Correlation of the entire proteome for analysis of structural and housekeeping functions is important. However, demonstrating the mimicry of the hepatocyte-specific proteome which governs its fundamental functions was of far greater relevance to this study. As such, proteins involved in metabolism or other hepatic functions were extracted from the dataset and compared.

5.2.3 Expression of Hepatic Markers and Proteins Involved in Metabolism

Pooled PHHs from 10 donors of varied age, gender, ethnicity, tobacco, alcohol and medication history (Lot number HUE50-D) were used as the hepatocyte "*gold standard*". Suspension hepatocytes for short term experiments (4 - 6 hours) with verified enzymatic activity are convenient for proteomic profiling, although the proteome generated is a snapshot of hepatocytes from donors following isolation and characterization. Despite being freshly isolated and cryopreserved to limit dedifferentiation, processing does reduced expression of several hepatic proteins, induce activation of pathways associated with cellular stress and augments the overall mature hepatic phenotype.

Metabolic properties associated with mature hepatocytes are thought to be governed by six core cross-regulating transcription factors. These factors are the hepatic nuclear factors, HNF1α, HNF1β, HNF3β, HNF4α1, HNF6 and LRH1 (Lemaigre F, 2010). Additionally, twelve genes with enriched expression in liver, and described as the liver specific proteome, have been identified in the human protein atlas (HPA). These are apolipoprotein A-II (P02652), alpha-1-B glycoprotein (P04217), alpha-2-HS-glycoprotein (P02765), coagulation factor II (P00451), complement factor H-related 2 (P36980), hemopexin (P369980), coagulation factor IX (P00740), secreted phosphoprotein 2 (Q13103), complement component 9 (P02748) mannose-binding lectin 2 (P1126) and CYP2A6 (P11509) (Uhlén M, Fagerberg L *et al.*, 2015).

Hepatic marker expression, reported in various publications (Baxter M, Withey S *et al.*, 2015; Lemaigre F, 2010; Schwartz R, Fleming H *et al.*, 2014; Si-Tayeb K, Noto FK *et al.*, 2010) and other findings was compiled and the relative quantitation of these identified marker proteins compared for the different hepatic cell types and culture techniques used. Of the 12 genes identified by the HPA, proteins associated with 6 genes were present in the filtered datasets collected in this study.

Additionally, plasma proteins, enzymes, bile proteins and transporters, present in, but not exclusive to the liver proteome, were also identified. These were summarise in Table 5.3 and although not an exhaustive list of proteins, the hepatic proteins extracted from the dataset demonstrate that PHH have higher relative protein abundance of the dominant markers. HLCs and HepG2 cells appear to have variable abundance of specific hepatic proteins whereby some proteins are more abundant in HLCs and others in HepG2 cells. Therefore, merely taking an overview of dominant hepatic proteins may not be sufficient in determining which model is more reflective of PHHs. There is however a trend of increased abundance of dominant liver marker proteins in HepG2 spheroids compared to the equivalent monolayers cultures.

Table 5.3: Hepatic marker proteins present in hepatic cells expressed as a ratio relative to PHH

Protein description	Hepatic cells				Main accession
	PHH	HLC	HepG2	HepG2_3D	
Alpha-1-acid glycoprotein 1	1.000	0.202	0.088	0.077	P02763
Alpha-1-antitrypsin	1.000	1.178	0.379	0.452	P01009
Alpha-1B-glycoprotein	1.000	0.164	0.147	0.365	P04217
Alpha-2-HS-glycoprotein	1.000	0.771	1.147	2.205	P02765
Alpha-fetoprotein	1.000	3.897	0.550	1.689	P02771
Amine oxidase [flavin-containing] A	1.000	0.141	0.082	0.189	P21397
Amine oxidase [flavin-containing] B	1.000	0.194	0.185	0.274	P27338
Angiotensinogen	1.000	1.377	0.418	0.752	P01019
Antithrombin-III	1.000	0.184	0.125	0.270	P01008
Apolipoprotein A-I	1.000	1.966	0.640	1.131	P02647
Apolipoprotein A-II	1.000	0.734	0.916	2.149	P02652
Apolipoprotein B-100	1.000	0.897	1.004	1.507	P04114
Apolipoprotein C-II	1.000	0.256	0.332	0.352	P02655
Asialoglycoprotein receptor 1	1.000	0.377	0.616	0.797	P07306
Asialoglycoprotein receptor 2	1.000	0.225	0.566	0.496	P07307
Bifunctional epoxide hydrolase 2	1.000	0.342	0.051	0.141	P34913
Bile salt export pump	1.000	0.170	0.106	0.126	O95342
Carbamoyl-phosphate synthase. mitochondrial	1.000	0.180	0.000	0.027	P31327
CCAAT/enhancer-binding protein alpha	1.000	0.418	1.025	1.378	P49715
Complement C5	1.000	0.535	0.352	0.344	P01031
Complement component C9	1.000	0.205	0.172	0.174	P02748
Complement factor B	1.000	0.405	0.207	0.311	P00751
C-reactive protein	1.000	0.065	0.013	0.045	P02741
Cytochrome P450 2A6	1.000	0.005	0.005	0.067	P11509
Dimethylaniline monooxygenase [N-oxide-forming] 5	1.000	0.089	0.102	0.203	P49326
Epoxide hydrolase 1	1.000	0.206	0.044	0.197	P07099
Fibrinogen alpha chain	1.000	0.684	0.195	0.916	P02671
Fibrinogen gamma chain	1.000	0.610	0.225	0.562	P02679
Glutathione peroxidase 1	1.000	0.364	0.404	0.279	P07203
Haptoglobin	1.000	0.255	0.000	0.032	P00738
Haptoglobin-related protein	1.000	0.292	0.007	0.052	P00739
Hemopexin	1.000	0.278	0.091	0.121	P02790
Heparin cofactor 2	1.000	0.759	0.492	1.008	P05546
Hepatocyte nuclear factor 3-alpha	1.000	0.000	0.000	0.000	P55317
Hepatocyte nuclear factor 4-alpha	1.000	0.497	1.182	1.778	P41235
Inter-alpha-trypsin inhibitor heavy chain H1	1.000	0.343	0.240	0.248	P19827
Inter-alpha-trypsin inhibitor heavy chain H2	1.000	0.178	1.303	1.602	P19823
Inter-alpha-trypsin inhibitor heavy chain H3	1.000	0.260	1.527	1.063	Q06033

Protein description	Hepatic cells				Main accession
	PHH	HLC	HepG2	HepG2_3D	
Inter-alpha-trypsin inhibitor heavy chain H4	1.000	0.090	0.720	0.248	Q14624
Keratin. type I cytoskeletal 18	1.000	0.522	0.632	0.639	P05783
Keratin. type II cytoskeletal 8	1.000	0.785	0.824	0.784	P05787
Mannose-binding protein C	1.000	0.164	0.076	0.043	P11226
Prospero homeobox protein 1	1.000	0.231	1.192	1.207	Q92786
Protein AMBP	1.000	0.631	0.752	0.933	P02760
Prothrombin	1.000	0.817	0.480	0.587	P00734
S-adenosylmethionine synthase isoform 1	1.000	0.525	0.058	0.308	Q00266
Serine--pyruvate aminotransferase	1.000	0.295	0.031	0.137	P21549
Serum albumin	1.000	0.383	0.177	0.519	P02768
Serum paraoxonase/arylesterase 1	1.000	0.000	0.000	0.005	P27169
Solute carrier organic anion transporter family member 1B1	1.000	0.056	0.032	0.058	Q9Y6L6
Solute carrier organic anion transporter family member 2B1	1.000	0.322	0.149	0.199	O94956
Transcription factor GATA-4	1.000	0.881	0.730	1.410	P43694
Transthyretin	1.000	1.336	0.100	0.431	P02766
Tyrosine aminotransferase	1.000	0.126	0.087	0.078	P17735
Vitronectin	1.000	0.739	0.309	0.508	P04004

Liver enriched transcription factors such as HNFs, in combination with CAAT enhancer binding proteins (CEBPs), regulate liver specific function. Target genes of HNF3 include transthyretin, albumin, AFP, APOB, APOA1, transferrin, A1AT, tyrosine aminotransferase, cholesterol-7α-hydroxylase, CYP2C6, phosphor-enol-pyruvate carboxykinase, phopshofructokinase 2 and aldolase B. Hepatoma cells generally express a non-functional HNF3α mutant which results in cessation in expression of many of these target genes (Selden C, Khalil M *et al.*, 1999).

Although HNF3α was identified in this study only PHHs showed levels that could be quantified. This could be due to the low abundance in the other cell types making detection of the reporter ions amongst the chromatographic noise disputable. Despite this notable absence of HNF3α, various CEBPs were present in HLCs and HepG2 cells along with many of the target proteins of this gene combination.

Metabolic capabilities required for sufficient hepatic responses after exposure to potential toxic compounds is based on the interplay of co-ordinated phase I, II and III enzymes. The nature and complexity of this organ system is reflected in the inability of models to accurately display detectable and quantifiable features essential in undertaking the well-defined functions of the liver.

When assessing whole proteome data, the overall phenotype may not exhibit changes in these metabolic capabilities merely due to the vast number of proteins being investigated or due to the fact that the fold changes is relatively minor. Changes in classical markers of the hepatic phenotype could be more or less differentially regulated than anticipated once extracted from the large data set. Data was interrogated specifically for proteins known to be involved in the various phases of hepatic metabolism (Table 5.4).

Table 5.4: Proteins involved in metabolism identified in hepatic cells expressed as a ratio relative to PHHs

Protein description	Hepatic cells				Main accession
	PHH	HLC	HepG2	HepG2_3D	
Alcohol dehydrogenase [NADP(+)]	1.000	0.542	0.243	0.339	P14550
Alcohol dehydrogenase 1A	1.000	0.000	0.000	0.049	P07327
Alcohol dehydrogenase 1B	1.000	0.011	0.000	0.058	P00325
Alcohol dehydrogenase 1C	1.000	0.052	0.019	0.072	P00326
Alcohol dehydrogenase 4	1.000	0.027	0.020	0.073	P08319
Alcohol dehydrogenase 6	1.000	0.036	0.066	0.211	P28332
Alcohol dehydrogenase class-3	1.000	0.850	0.504	0.687	P11766
Aldehyde dehydrogenase family 16 member A1	1.000	0.561	0.626	0.718	Q8IZ83
Aldehyde dehydrogenase family 8 member A1	1.000	0.084	0.028	0.015	Q9H2A2
Aldehyde dehydrogenase X, mitochondrial	1.000	0.077	0.393	0.465	P30837
Aldehyde dehydrogenase, mitochondrial	1.000	0.177	0.143	0.194	P05091
Cytochrome b5	1.000	0.366	0.051	0.334	P00167
Cytochrome P450 1A1	1.000	0.416	0.118	0.333	P04798
Cytochrome P450 1A2	1.000	0.000	0.108	0.061	P05177
Cytochrome P450 20A1	1.000	0.484	0.573	0.589	Q6UW02
Cytochrome P450 2A6	1.000	0.005	0.005	0.067	P11509
Cytochrome P450 2B6	1.000	0.113	0.066	0.071	P20813
Cytochrome P450 2C19	1.000	0.143	0.031	0.048	P33261
Cytochrome P450 2C8	1.000	0.079	0.008	0.074	P10632
Cytochrome P450 2C9	1.000	0.084	0.000	0.070	P11712
Cytochrome P450 2D6	1.000	0.049	0.017	0.027	P10635
Cytochrome P450 2E1	1.000	0.175	0.058	0.143	P05181
Cytochrome P450 2J2	1.000	0.152	0.092	0.103	P51589
Cytochrome P450 2S1	1.000	7.980	3.025	2.197	Q96SQ9
Cytochrome P450 2W1	1.000	0.649	1.463	2.214	Q8TAV3
Cytochrome P450 3A4	1.000	0.115	0.000	0.041	P08684
Cytochrome P450 3A5	1.000	0.076	0.052	0.070	P20815

Protein description	Hepatic cells				Main accession
	PHH	HLC	HepG2	HepG2_3D	
Cytochrome P450 4A11	1.000	0.064	0.000	0.000	Q02928
Cytochrome P450 4V2	1.000	0.159	0.068	0.067	Q6ZWL3
Dimethylaniline monooxygenase [N-oxide-forming] 3	1.000	0.042	0.032	0.098	P31513
Dimethylaniline monooxygenase [N-oxide-forming] 4	1.000	0.165	0.127	0.134	P31512
Dimethylaniline monooxygenase [N-oxide-forming] 5	1.000	0.089	0.102	0.203	P49326
Gamma-glutamyltranspeptidase 1	1.000	2.224	0.572	1.101	P19440
Glutathione S-transferase A2	1.000	0.749	0.022	0.254	P09210
Glutathione S-transferase C-terminal domain- protein	1.000	1.078	1.571	1.375	Q8NEC7
Glutathione S-transferase kappa 1	1.000	0.491	0.204	0.290	Q9Y2Q3
Glutathione S-transferase Mu 1	1.000	0.097	0.080	0.095	P09488
Glutathione S-transferase Mu 2	1.000	0.980	0.539	0.756	P28161
Glutathione S-transferase Mu 3	1.000	2.254	1.357	1.302	P21266
Glutathione S-transferase omega-1	1.000	0.774	0.313	0.499	P78417
Glutathione S-transferase P	1.000	8.621	0.000	0.136	P09211
Glutathione S-transferase theta-1	1.000	0.263	0.502	0.722	P30711
Glutathione S-transferase theta-2	1.000	0.220	0.106	0.092	P0CG29
Multidrug resistance protein 1	1.000	0.329	1.328	0.936	P08183
Multidrug resistance-associated protein 1	1.000	1.925	3.272	2.014	P33527
Multidrug resistance-associated protein 6	1.000	0.171	0.223	0.324	O95255
NADPH-CYP450 reductase	1.000	0.518	0.403	0.556	P16435
UDP-glucuronosyltransferase 1-1	1.000	0.195	0.000	0.067	P22309
UDP-glucuronosyltransferase 1-4	1.000	0.183	0.000	0.090	P22310
UDP-glucuronosyltransferase 1-6	1.000	0.151	0.017	0.047	P19224
UDP-glucuronosyltransferase 2A3	1.000	0.334	0.548	0.794	Q6UWM9
UDP-glucuronosyltransferase 2B10	1.000	0.047	0.011	0.143	P36537
UDP-glucuronosyltransferase 2B15	1.000	0.067	0.071	0.086	P54855
UDP-glucuronosyltransferase 2B17	1.000	0.087	0.044	0.071	O75795
UDP-glucuronosyltransferase 2B4	1.000	0.010	0.000	0.055	P06133
UDP-glucuronosyltransferase 2B7	1.000	0.078	0.028	0.137	P16662

Phase I, II and III enzymes encompassing CYP450s, monoamine oxidases, flavin-containing oxygenases, esterases, UGTs, sulfotransferases, N-acetyl transferases, glutathione S-transferases and multi-drug resistance proteins were identified. As would be expected, PHHs expression of these enzymes was highest in comparison to the other cell types which all demonstrated pronounced variability.

Copies of mRNA in HepG2 cells exhibit a shortage of phase I enzymes (CYP1A1, CYP1A2, CYP2B6, CYP2C9, CYP2D6, CYP2E1 and CYP3A4) often limiting the use of HepG2 cells. CYP3A7, the foetal counterpart closely related to CYP3A4, contains more mRNA copies in HepG2 cells than PHHs. Wilkening *et al.* suggest that this, along with some other factors, shows HepG2 cells resemble a foetal phenotype. In contrast, some phase II enzymes of drug metabolism, can be more highly expressed in HepG2 cells (EPHX1, GSTM1, UGT1A6)(Wilkening S, Stahl F *et al.*, 2003).

Of the phase II enzymes assessed by Wilkening *et al.* (EPHX1, GSTM1, UGT1A6, SULT1A1, NQO1, NAT1), epoxide hydrolase 1 (P07099), glutathione S-transferase Mu 1 (P09488), UDP-glucuronosyltransferase 1-6 (P19224), NAD(P)H dehydrogenase [quinone] 1 (P15559) and arylamine N-acetyltransferase 1 (P18440) were identified in the proteomic data. However of these, only NQO1 had higher expression in HepG2 cells illustrating a discordance between mRNA and protein expression for the other proteins.

What is apparent is that HepG2 spheroids appear to have enhanced metabolic capabilities over HepG2 cell monolayers despite not rivalling PHHs. Inconsistent expression trends for metabolic proteins in HLCs make determination of the more applicable model difficult. The low levels of metabolic proteins in HLCs could be attributed to various factors. Firstly, the heterogeneous nature of HLC differentiation may result in larger of variance if the degree of differentiation is not tightly regulated. Secondly, as seen in Chapter 4, the hepatic proteome of HLCs appears to decline between Day 30 and Day 32. Comparing PHHs to HLCs collected after 35 days of differentiation could be grossly underestimating the hepatic phenotype of this cell type, purely based on the time point of collection.

In addition to observations of known markers of hepatic phenotypes and maturation, reference protein profiles (Figure 5.11) were generated for proteins which were elevated in a specific cell type but maintained linearity in abundance for each of the other three models. Using these profiles, the top 20 proteins which cluster together were identified and Uniprot used to annotate the molecular or biological functions (Table 5.5) in an to attempt to ascertain any major distinguishing features of each model.

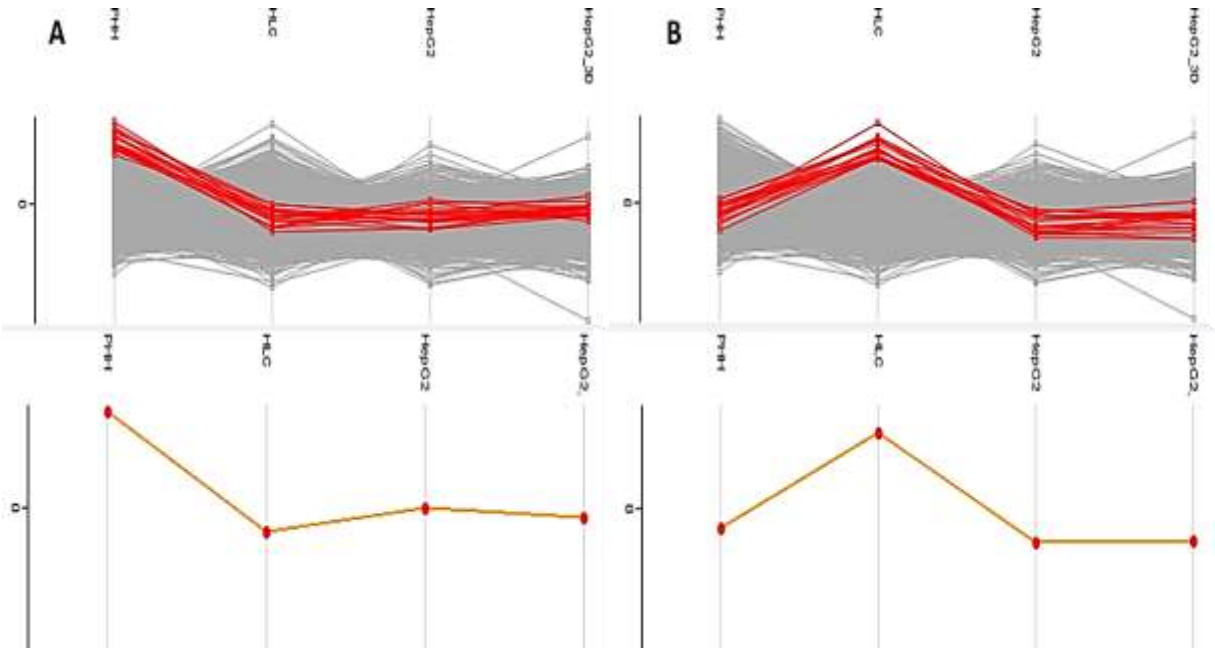

Figure 5.11: Protein profiles (top) and reference profiles (bottom) generated to cluster the top 20 proteins which peak in A) PHHs and B) HLCs. Labels from left to right: PHH, HLC, HepG2 and HepG2_3D.

Table 5.5: Proteins peaking in specific cell models but further maintained linearity in expression (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in PHHs		
Perilipin-4	Q96Q06	Triacylglycerol packaging into adipocytes Coat protein involved in biogenesis of lipid droplets
Solute carrier family 22 member 7	Q9Y694	Active cellular uptake of carnitine Transports of organic cations
Regulator of microtubule dynamics protein 2	Q96LZ7	Interphase localization in the cytoplasm Spindle microtubules and poles localization during mitosis
Xanthine dehydrogenase/oxidase	P47989	Key enzyme in purine degradation Catalyse oxidation of xanthine to uric acid Generation of reactive oxygen species
Dimethylaniline monooxygenase [N-oxide-forming] 4	P31512	Oxidative metabolism of a variety of xenobiotics Microsome and endoplasmic reticulum localization
Alpha-aminoadipic semialdehyde synthase, mitochondrial	Q9UDR5	Bifunctional enzyme Catalyses the first two steps in lysine degradation
Mitochondrial glutamate carrier 2	Q9H1K4	Transport of glutamate across the inner mitochondrial membrane
Fatty aldehyde dehydrogenase	P51648	Oxidation of long-chain aliphatic aldehydes to fatty acids
C4b-binding protein alpha chain	P04003	Controls classical pathway of complement activation Binds as a cofactor to C3b/C4b inactivator
Acyl-CoA-binding domain-containing protein 5	Q5T8D3	Integral membrane protein of the peroxisome Peroxisome receptor for pexophagy

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in HLCs		
Kunitz-type protease inhibitor 2	O43291	Serine protease inhibitor Hepatocyte growth factor activator inhibitor type 2 Inhibits plasmin, plasma and tissue kallikrein, and factor XIa
Heat shock protein beta-8	Q9UJY1	Temperature-dependent chaperone activity
Fatty acid-binding protein, heart	P05413	Intracellular transport of long-chain fatty acids and acyl-CoA esters
Podocalyxin	O00592	Regulation of both adhesion and cell morphology Initial epithelial polarization and apical lumen formation
Transmembrane channel-like protein 4	Q7Z404	Probable ion channel
Sialate O-acetylesterase	Q9HAT2	Removal of O-acetyl ester groups from sialic acid Carbohydrate metabolic processes
CD63 antigen	P08962	Activation of cellular signalling cascades: AKT and FAK/PTK2 Cell survival, reorganization of actin cytoskeleton, cell adhesion, spreading and migration
Lactadherin	Q08431	Binds to phosphatidylserine-enriched cell surfaces Phagocytic removal of apoptotic cells
Matrix metalloproteinase-14	P50281	Potential actin cytoskeleton reorganization Positive regulator of cell growth and migration
Adenosine deaminase	P00813	Hydrolytic deamination of adenosine Purine metabolism and in adenosine homeostasis
Proteins peaking in HepG2 monolayers		
Complexin-2	Q6PUV4	Positively regulates a late step in exocytosis of various cytoplasmic vesicles
Keratin, type I cytoskeletal 23	Q9C075	Structural molecule
Neural cell adhesion molecule L1	P32004	Developmental protein involved in cell adhesion
Cystine/glutamate transporter	Q9UPY5	Sodium-independent, high-affinity exchange of anionic amino acids
Cell surface glycoprotein MUC18	P43121	Cell adhesion and cohesion of the endothelial monolayer at intercellular junctions
Semaphorin-3B	Q13214	Semaphorin family implicated on promoting or inhibiting tumour progression
Cartilage oligomeric matrix protein	P49747	Role in the structural integrity Suppresses apoptosis in transformed cells
Alpha-catulin	Q9UBT7	Modulates the Rho pathway signalling
Protein sprouty homolog 4	Q9C004	Negative regulation of the ERK1 and ERK2 cascades
PDZ and LIM domain protein 7	Q9NR12	May function as a scaffold for coordinated protein assembly BMP6 signalling pathway
Proteins peaking in HepG2 spheroids		
N-acetylated-alpha-linked acidic dipeptidase-like protein	Q9UQQ1	Release of unsubstituted, C-terminal glutamyl residue Dipeptidyl-peptidase IV type activity
Apolipoprotein A-II	P02652	Affect high density lipoprotein (HDL) metabolism May stabilize HDL structure
Basal cell adhesion molecule	P50895	May mediate intracellular signalling Laminin alpha-5 receptor Major component of basement membrane

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in HepG2 spheroids		
Immunoglobulin superfamily member 1	Q8N6C5	Co-receptor in inhibin signalling Antagonizes activin A signalling
Protein phosphatase 1H	Q9ULR3	Dephosphorylates CDKN1B Removing a signal for proteasomal degradation
Espin	B1AK53	Multifunctional actin-bundling protein Regulating organization, dynamics and signalling capacities
Centrin-2	P41208	Microtubule organizing centre structure and function
Dual specificity mitogen-activated protein kinase kinase3	P46734	Catalyzes the concomitant phosphorylation of a threonine and a tyrosine residue in MAP kinase p38
Rho guanine nucleotide exchange factor 40	Q8TER5	Guanine nucleotide exchange factor Regulation of Rho signalling
Leucine-rich repeat-containing protein 41	Q15345	Probable substrate recognition component of an ECS E3 ubiquitin ligase complex

Overviews of proteins reduced in a cell type but maintaining linearity in abundance for the other models were not explicitly linked to a molecular or biological function. When investigating such large comparative data sets it is important to consider not only the implications of which proteins are identified and their relative abundance, but also which proteins are absent. Absence could be due to the inability of the instrumentation to detect low abundance proteins, filtering out of proteins during analysis or merely an absence of the protein within a sample.

5.2.4 Cell Type Comparisons: PHHs versus HLCs

Differentiation of iPSCs to HLCs requires a complex interplay of cell proliferation and maturation. Numerous differentiation protocols show conversion into hepatic lineages which produce HLCs of varying maturity. HLCs should therefore be benchmarked against iPSCs and both adult and foetal hepatocytes to establish where the HLCs fall within the continuum of maturation. Foetal hepatocytes for experimental purposes remains limited and hepatic tissue for safety pharmacology should mimic PHHs. Furthermore, rapid dedifferentiation of PHH used as the "*gold standard*" may result in over-interpretation of hepatic cell phenotypes (Baxter M, Withey S *et al.*, 2015). Based on the abovementioned factors, HLCs were only compared to freshly isolated and cryopreserved PHHs.

The use of volcano plots for easy identification of changes in large comparative datasets is commonly used in proteomic analysis. The relative importance of the t-test value and the difference between the mean can be adjusted in relation to the data set size. Volcano plots were generated using a t-test with 250 randomizations at an FDR of 0.01 with controlling only for the p-value (s0 = 0). Additionally, incorporating a relative weighting of the mean differences (s0 = 0.1) at the same FDR (Figure 5.12) was done and the plots used for downstream assessments (Tusher VG, Tibshirani R *et al.*, 2001).

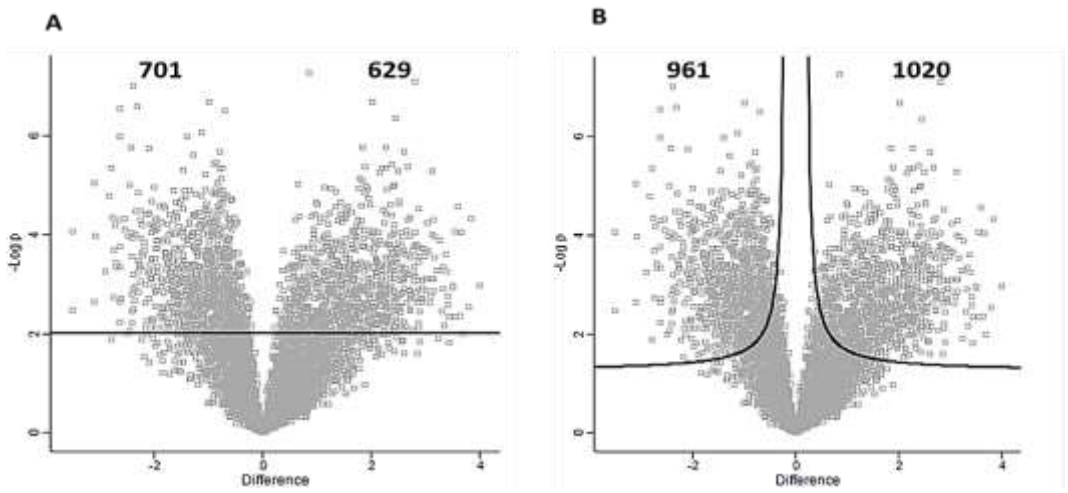

Figure 5.12: Volcano plots comparing PHHs to HLCs using an FDR of 0.01 with A) Contribution from the p-value only and B) Contributions from the p-value and mean difference. Values above and to the left or right were differentially down- or up-regulated relative to PHHs respectively.

Interpretation of volcano plots needs to be done with care. Proteins which show a positive difference on the x-axis means that the protein had greater expression in the first grouping. Therefore, in the volcano plot above, proteins on the right were proteins with higher in PHHs versus HLCs whereas proteins on the left were proteins higher in HLCs versus PHHs. Comparing PHHs to HLCs with inclusion of the difference in the mean (Figure 5.12B), illustrated 961 proteins increased and 1 020 decreased in HLCs. Some outliers (Figure 5.13) on the volcano plots were manually selected and annotated for function (Table 5.6).

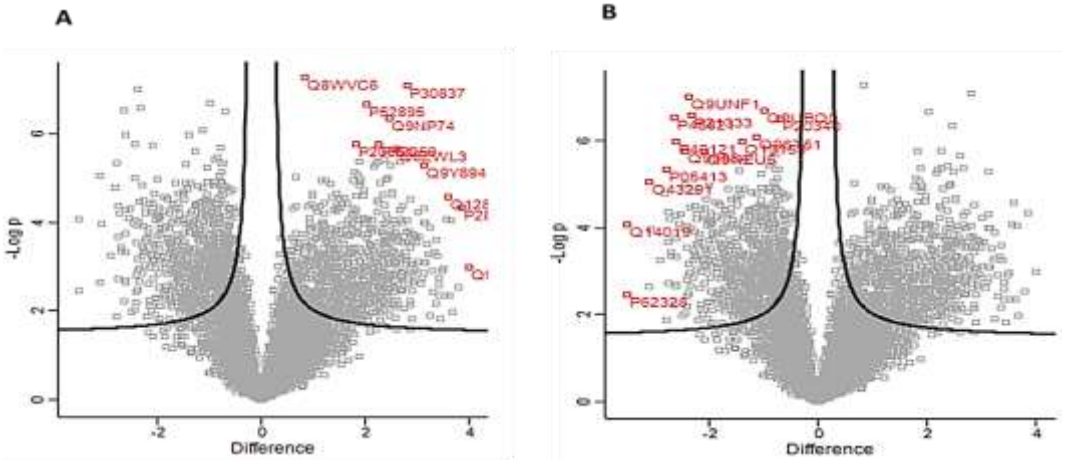

Figure 5.13: Proteome comparison of PHHs versus HLCs using volcano plots. Proteins located at either the extreme right or extreme left of volcano plots of PHHs versus HLCs which were investigated and described.

Proteins with high significance values or differences for PHHs (Table 5.6) included proteins involved in fatty acid and lipid synthesis and metabolism (Q8WVC6, P52895, Q6ZWL3) as well as important detoxifiers aldehyde (P30837) and alcohol dehydrogenase (P28332). Actins are highly conserved ubiquitously expressed proteins, involved in various cellular processes including cell motility, cell division, cytokinesis as well as the establishment and maintenance of cell morphology (TheUniProtConsortium, 2014).

Actin-associated proteins (Q14019, P62328, P46821, P21333) were increased in HLCs compared to PHHs. Despite having morphological similarities to plated PHHs, these proteins suggest differences in the cytoskeletal arrangement between the two cell types. This could be due to the distinct morphological changes which HLCs undergo during differentiation and the nature of culturing HLCs as confluent monolayers for differentiation. Additionally, Kunitz-type protease inhibitor 2 (O43291), an inhibitor of hepatocyte growth factor activator, was increased in HLCs assayed after 30 days of differentiation. The implications of increased inhibition of HGF activation could have consequences for hepatocyte maturity, proliferation, adhesion and apoptosis.

Table 5.6: Proteins with large differential abundance between PHHs versus HLCs
(TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with higher expression in PHHs compared to HLCs				
Dephospho-CoA kinase domain-containing protein	Q8WVC6	Biosynthesis of co-enzyme A for synthesis and oxidation of fatty acids	7.272	0.857
C4b-binding protein beta chain	P20851	Classical pathway of complement activation	5.761	1.840
Aldo-keto reductase family 1 member C2	P52895	Conversion of steroid hormones Lipid and steroid metabolism	6.667	2.033
Coagulation factor V	P12259	Central regulator of haemostasis Cofactor for prothrombinase activity	5.770	2.274
Palmdelphin	Q9NP74	Regulation of cell shape	6.350	2.458
Cytochrome P450 4V2	Q6ZWL3	Oxidize medium-chain saturated fatty acids	5.668	2.600
Aldehyde dehydrogenase X, mitochondrial	P30837	Detoxification of alcohol-derived acetaldehyde, metabolism of corticosteroids and biogenic amines	7.094	2.819
Solute carrier family 22 member 7	Q9Y694	Sodium-independent multi-specific organic anion transport	5.293	3.125
Nuclear factor 1 A-type	Q12857	Activating transcription and replication	4.573	3.587
Alcohol dehydrogenase 6	P28332	Alcohol dehydrogenase involved in ethanol oxidation	4.325	3.852
Dimethylglycine dehydrogenase, mitochondrial	Q9UI17	Sub-pathway that synthesizes sarcosine (by-product in glycine synthesis and degradation)	2.978	4.007

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with lower expression in PHHs compared to HLCs				
Coactosin-like protein	Q14019	Calcium-independent binding to F-actin	4.070	-3.512
Thymosin beta-4	P62328	Organization of the cytoskeleton Inhibits actin polymerization	2.473	-3.498
Kunitz-type protease inhibitor 2	O43291	Inhibitor of hepatocyte growth factor (HGF) activator	5.042	-3.104
Fatty acid-binding protein. heart	P05413	Intracellular transport of long-chain fatty acids	5.344	-2.795
Microtubule-associated protein 1B	P46821	Tyrosination of alpha-tubulin	6.537	-2.635
Macrophage-capping protein	P40121	Reversibly blocks the barbed ends of actin filaments	5.984	-2.621
Protein S100-A13	Q99584	Export of proteins that lack a signal peptide	5.754	-2.425
Melanoma-associated antigen D2	Q9UNF1	Hepatocellular carcinoma-associated protein. widely expressed in normal tissue	6.996	-2.373
Filamin-A	P21333	Orthogonal branching of actin filaments	6.584	-2.325
LIM and cysteine-rich domains protein 1	Q9NZU5	Transcriptional cofactor that restricts GATA6 function	5.741	-2.093
Serine/threonine-protein kinase PAK 1	Q13153	Role in cytoskeleton dynamics, cell adhesion, migration, proliferation, apoptosis and mitosis	5.970	-1.383
RUN and FYVE domain-containing protein 1	Q96T51	Binds phospholipid vesicles containing phosphatidylinositol 3-phosphate	6.065	-1.120
Vacuolar protein sorting-associated protein 29	Q9UBQ0	Retromer cargo-selective complex Prevent missorting of transmembrane cargo proteins into lysosomal degradation	6.682	-0.984
Ras-related protein Rab-6A	P20340	Regulator of membrane traffic from the Golgi apparatus towards the ER	6.501	-0.698

Differentiation of iPSCs can be associated with variation in differentiation efficiency and therefore varied maturation. As a result, a minimum of two different differentiations was included per 6-plex experiment to account for potential variability. Based on the time course of HLC differentiation, proteome level data suggests that differentiation after 30 days results in a reduced hepatic function. Therefore, underestimation of the proteomic correlation between iPSC-HLCs to PHHs is possible. Assessment of several additional hepatic functions such as glycogen synthesis, lipid accumulation and urea metabolism may have aided in determining the optimal time point for termination of differentiation. The ability to definitively establish whether HLCs could rival PHHs in toxicity testing would make use of more comprehensive data from the iPSCs to HLC time course to provide information for better decision making on sample collection and also compare multiple maturation endpoints simultaneously with PHHs. This proteomic data should also be matched to toxicity testing outcomes initiated at the same time points that the proteome is evaluated.

5.2.5 Cell Type Comparisons: PHHs versus HepG2 Monolayers

HepG2 cells have been associated with discontinuous phenotypic and functional properties (Duret C, Gerbal-Chaloin S *et al.*, 2007). Additionally, gene expression levels, notably of phase I drug metabolizing enzymes, in PHHs differ significantly from those in HepG2 cells (Wilkening S, Stahl F *et al.*, 2003). Comparing a discontinuous reduced HepG2 phenotype, which can become altered over various sequential passages, to PHHs which express a phenotype that is lot and isolation dependent can influence the biological implications of the data described.

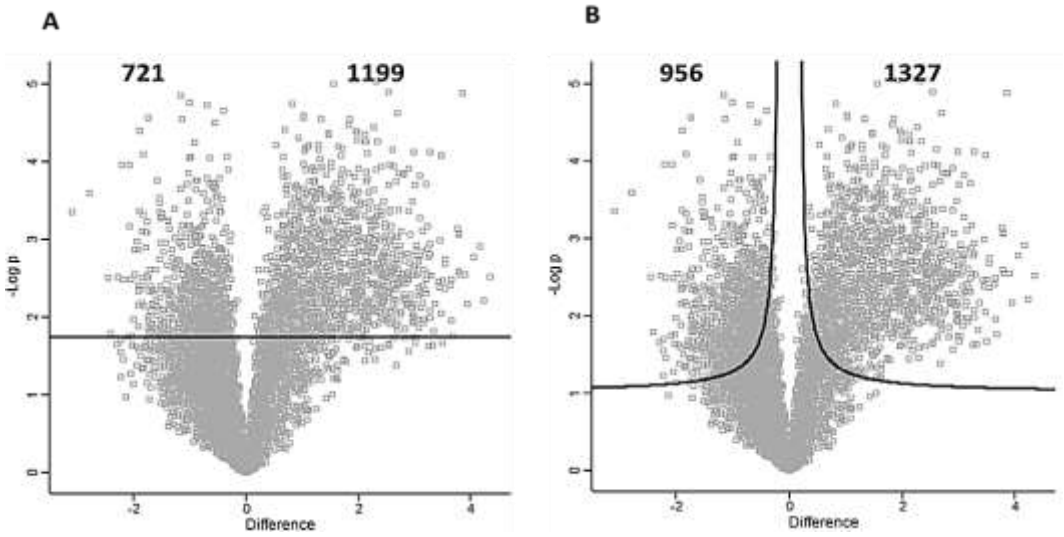

Figure 5.14: Volcano plots comparing PHHs to HepG2 monolayers using an FDR of 0.01 with A) Contribution from the p-value only and B) Contributions from the p-value and mean difference. Values above and to the left or right were differentially down- or up-regulated relative to PHHs respectively.

Comparing PHHs to HepG2 monolayers (Figure 5.14B) demonstrated 956 proteins which were increased and 1 327 which were decreased in HepG2 monolayers. Based on filtering for presence in all three replicates, 5 231 proteins were analysed. This means that the differentially regulated proteins, whether increased or decreased, accounted for 43.64% of proteins identified and quantified. Some outliers (Figure 5.15) were manually selected and annotated for function (Table 5.7) to investigate the potential implicates of these highly differentially expressed proteins.

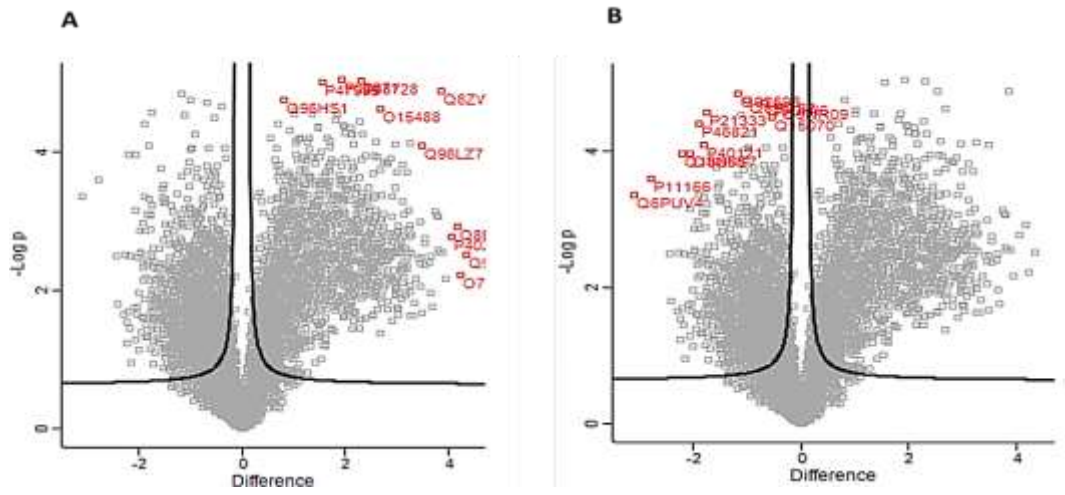

Figure 5.15: Proteome comparison of PHHs versus HepG2 monolayers using volcano plots. Proteins located at either the extreme right or extreme left of volcano plots of PHHs versus HepG2 monolayers which were investigated and described.

Proteins with high significance values or expression in PHHs included many mitochondrial proteins involved in energy production and catalysis (Table 5.7). Metabolic proteins involved in xenobiotic metabolism (P40261), fatty acid metabolism (Q6ZWL3) and transport or clearance (O76054, Q9Y6L6), which are essential to facilitate integral hepatic functions were notably absent in HepG2 cells.

Actin-associated proteins (Q16658; Q8N8S7; P46821; P40121; P21333) were identified as having high expression in HepG2 monolayers. These proteins function as anchors, scaffolds and signalling proteins in all cells. However, being related to the formation of filopodia and playing a role in cell migration could explain this high level of expression compared to isolated PHHs cultured in suspension.

Cellular morphology can dictate functional characteristics. Isolated PHH are typically cuboid in shape with potential multinucleation, whereas HepG2 cells display an epithelial-like morphology with a single nucleus (Wilkening S, Stahl F *et al.*, 2003). HepG2 cells cultured on artificial substrates in monolayers display cell protrusions for motility and migration not seen *in vivo* or in suspensions of PHHs. This unnatural cell culture microenvironment could be responsible for this enhance expression of multiple ubiquitous actin-associated proteins.

Table 5.7: Proteins with large differential abundance between PHHs versus HepG2 monolayers
(TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with higher expression in PHHs compared to HepG2 monolayers				
Serine/threonine-protein phosphatase PGAM5, mitochondrial	Q96HS1	Dephosphorylates and activates MAP3K5 kinase	4.748	0.822
Cytochrome b-c1 complex subunit Rieske, mitochondrial	P47985	Respiratory chain that generates an electrochemical potential coupled to ATP synthesis	5.003	1.563
Fibrinogen alpha chain	P02671	Polymerize to form an insoluble fibrin matrix	5.049	1.941
Aminomethyltransferase, mitochondrial	P48728	Catalyses the degradation of glycine	5.025	2.321
Glycogenin-2	O15488	Form an oligosaccharide primer that serves as substrate for glycogen synthase	4.621	2.692
Regulator of microtubule dynamics protein 2	Q96LZ7	Localizes to spindle microtubules and spindle poles during mitosis	4.083	3.483
Cytochrome P450 4V2	Q6ZWL3	Fatty acid and steroid metabolism	4.878	3.857
Nicotinamide N-methyltransferase	P40261	N-methylation of nicotinamide to form pyridinium ions Biotransformation of xenobiotics	2.768	4.066
Methylmalonic aciduria type A protein, mitochondrial	Q8IVH4	Probable GTPase May function as chaperone.	2.917	4.172
SEC14-like protein 2	O76054	Carrier protein of some hydrophobic molecules	2.220	4.234
Solute carrier organic anion transporter family member 1B1	Q9Y6L6	Na^+-independent uptake of organic anions Clearance of bile acids and organic anions from the liver	2.513	4.353
Proteins with lower expression in PHHs compared to HepG2 monolayers				
Complexin-2	Q6PUV4	Negatively regulates the formation of synaptic vesicle clustering	3.350	-3.099
Solute carrier family 2, facilitated glucose transporter member 1	P11166	Facilitative glucose transporter isoform may be responsible for basal glucose uptake	3.596	-2.790
Fascin	Q16658	Organizes filamentous actin into bundles	3.961	-2.196
Protein enabled homolog	Q8N8S7	Actin-associated proteins, cytoskeleton remodelling and cell polarity	3.960	-2.064
Microtubule-associated protein 1B	P46821	Tyrosination of alpha-tubulin	4.400	-1.882
Macrophage-capping protein	P40121	Reversibly blocks the barbed ends of actin filaments	4.085	-1.813
Filamin-A	P21333	Orthogonal branching of actin filaments and links actin filaments to membrane glycoproteins	4.566	-1.732

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with lower expression in PHHs compared to HepG2 monolayers				
Heat shock protein 105 kDa	Q92598	Prevents the aggregation of denatured proteins in cells under severe stress	4.850	-1.160
UHRF1-binding protein 1	Q6BDS2	Negative regulator of cell growth	4.760	-0.993
Probable ATP-dependent RNA helicase DDX20	Q9UHI6	Assembly of small nuclear ribonucleoproteins (building blocks of the spliceosome)	4.724	-0.691
Mitochondrial inner membrane protein OXA1L	Q15070	Insertion of integral membrane proteins into the mitochondrial inner membrane	4.497	-0.550
Baculoviral IAP repeat-containing protein 6	Q9NR09	Anti-apoptotic protein which can regulate cell death by controlling caspases	4.657	-0.394

5.2.6 Cell Type Comparisons: PHHs versus HepG2 Spheroids

Hanging drops as a means of 3D culture is easy to adopt and allows for the establishment of spatial organisation with cell-cell interactions. This enhances or simulates physiological boundaries more closely than monolayer cultures (Fey SJ and Wrzesinski K, 2012). Data in support of the extent to which 3D cultures can be used to minimize extrapolation and mimic *in vivo* counterparts is however fragmentary. Comparing the proteomes of PHHs to HepG2 spheroids (Figure 5.16B) identified 180 proteins which were increased and 682 which were decreased in HepG2 spheroids. This total variance accounts for only 18.8% (180 versus 956) and 51.4% (682 versus 1327) of the differences seen in HepG2 monolayers. Details and additional experiments comparing monolayers and spheroids are discussed in Chapter 7. However, it is evident that HepG2 spheroids more closely resemble PHHs than HepG2 cells under the experimental conditions described in this study. Some outliers (Figure 5.17) were manually selected and annotated for function (Table 5.8).

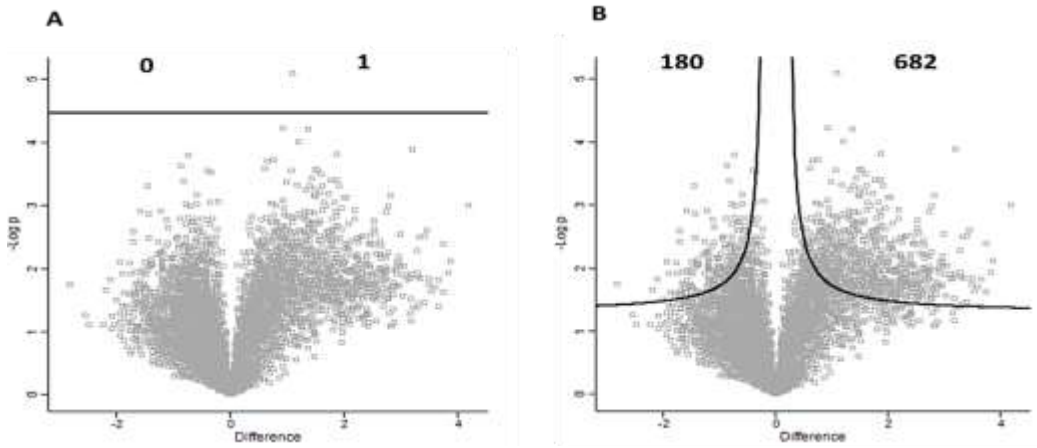

Figure 5.16: Volcano plots comparing PHHs to HepG2 spheroids using an FDR of 0.01 with A) Contribution from the p-value only and B) Contributions from the p-value and mean difference. Values to the left or right were differentially down- or up-regulated relative to PHHs respectively.

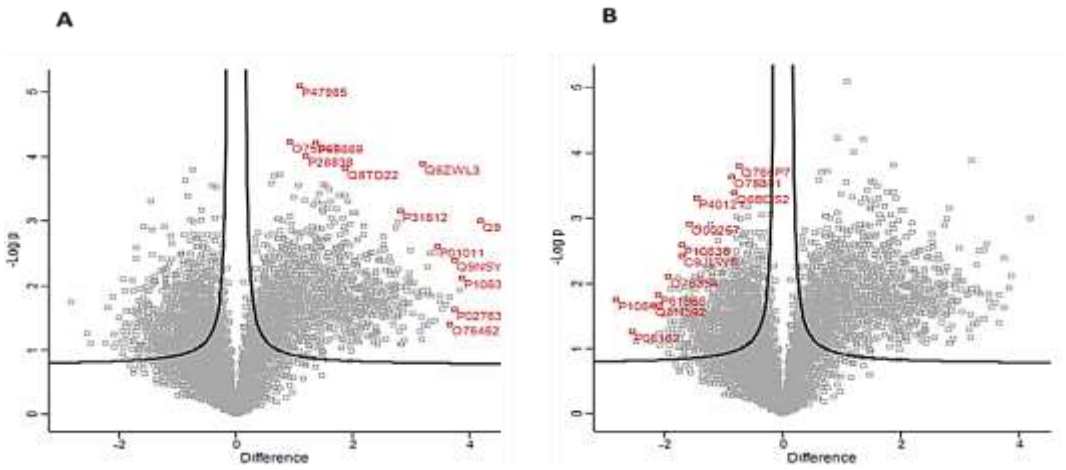

Figure 5.17: Proteome comparison of PHHs versus HepG2 spheroids using volcano plots. Proteins located at either the extreme right or extreme left of volcano plots of PHHs versus HepG2 spheroids which were investigated and described.

Similar to HepG2 monolayers, proteins with high significance values or differences compared to PHHs (Table 5.8) included proteins involved in energy production, transport and metabolism. These proteins, mainly mitochondrial in origin, are strongly associated with common hepatic functions. Furthermore, when the difference in the mean is not considered in producing volcano plots (Figure 5.16A), only one protein is considered significant. This outlier, also seen in monolayers, is a member of the cytochrome b-c1 complex (P47985) found in the inner mitochondrial membrane, which generates an electrochemical potential coupled to ATP synthesis.

Despite the ability to reduce the number of differentially regulated proteins by culturing in 3D, the model remains unable to capture some critical hepatic features. The absence of phase I and phase II drug metabolizing enzymes remains concerning. The main metabolizing CYP isoforms (1A2, 2C9, 2C19, 2D6 and 3A4) accounting for approximately 75% of all drug metabolism were present in PHHs. These six isoforms were poorly quantified in HepG2 monolayers with culturing as spheroids unable to salvage the lack of abundance in expression of these proteins. This could suggest that the 3D cultures are augmenting proteins involved in morphology, cell survival and other homeostatic associated features as opposed to hepatic maturity.

Table 5.8: Proteins with large differential abundance between PHHs versus HepG2 spheroids
(TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with higher expression in PHHs compared to HepG2 spheroids				
ATP synthase subunit, mitochondrial	O75947	Produces ATP from ADP in the presence of a proton gradient	4.229	0.916
Cytochrome b-c1 complex subunit Rieske, mitochondrial	P47985	Respiratory chain that generates an electrochemical potential coupled to ATP synthesis	5.102	1.082
Cytosol aminopeptidase	P28838	Processing and regular turnover of intracellular proteins	4.017	1.189
Cytochrome c oxidase subunit 6C	P09669	Terminal oxidase in mitochondrial electron transport	4.204	1.371
Sideroflexin-5	Q8TD22	Transmembrane citrate transport	3.819	1.859
Dimethylaniline monooxygenase [N-oxide] 4	P31512	Oxidative metabolism of xenobiotics	3.151	2.809
Cytochrome P450 4V2	Q6ZWL3	Fatty acid and steroid metabolism	3.894	3.193
Alpha-1-antichymotrypsin	P01011	Physiological function is unclear, can inhibit activation of angiotensin-1	2.609	3.439
Retinol dehydrogenase 16	O75452	Oxidoreductase (preference for NAD)	1.389	3.649
Alpha-1-acid glycoprotein 1	P02763	Transport protein in the blood stream	1.621	3.737
StAR-related lipid transfer protein 5	Q9NSY2	Intracellular transport of sterols or other lipids	2.389	3.746
Cytochrome P450 2D6	P10635	Metabolism of many drugs and environmental chemicals that it oxidizes	2.112	3.865
Oxidoreductase HTATIP2	Q9BUP3	Oxidoreductase for tumour suppression	3.001	4.188
Proteins with lower expression in PHHs compared to HepG2 spheroids				
Tissue factor pathway inhibitor	P10646	Antithrombotic action, inhibits factor X and VIIa/tissue factor activity	1.744	-2.833
Galectin-2	P05162	Protein binds beta-galactoside	1.254	-2.544
Rho GTPase-activating protein 18	Q8N392	Suppresses F-actin polymerization by inhibiting Rho	1.660	-2.196
Small ubiquitin-related modifier 2	P61956	Nuclear transport, DNA replication and repair, mitosis and signal transduction	1.818	-2.108
Ectonucleoside triphosphate diphosphohydrolase 6	O75354	Glycosylation reactions in the Golgi apparatus	2.093	-1.938
Microtubule-associated protein tau	P10636	Microtubule assembly and stability	2.595	-1.712
MAPK-regulated corepressor-interacting protein 1	C9JLW8	Molecular switch to regulate epithelial-mesenchymal transition	2.421	-1.702
E3 SUMO-protein ligase CBX4	O00257	Indirectly regulates p53/TP53 transcriptional activation	2.913	-1.580
Macrophage-capping protein	P40121	Reversibly blocks the barbed ends of actin filaments	3.312	-1.460
Sperm-associated antigen 7	O75391	Nucleic acid binding	3.629	-0.868
UHRF1-binding protein 1	Q6BDS2	Negative regulator of cell growth	3.382	-0.829
MTSS1-like protein	Q765P7	Function in actin bundling	3.792	-0.734

5.2.7 Comparison of Protein Expression

When comparing PHHs to HLCs, HepG2 monolayers and HepG2 spheroids, a larger number of proteins were found to be differentially regulated. PHH compared to HLCs, HepG2 monolayers and HepG2 spheroids demonstrated 1 981, 2 283 and 862 proteins respectively which were up- or down-regulated compared to PHHs. This posed an important series of questions: Are the same proteins up- or down-regulated in HLCs, HepG2 monolayers and HepG2 spheroids? Are some proteins up-regulated in one cell model yet down-regulated in another and vice versa? Can certain proteins (or protein function) aid in describing the large number of differentially regulated proteins?

The data generated from volcano plots was interrogated to determine the similarity and variability in protein expression between HLCs, HepG2 monolayers and HepG2 spheroids. This provided insight into whether all three models lacked expression of a similar cohort of proteins.

Table 5.9: Comparison of differentially expressed proteins were compared to establish similarity in positive, negative and multidirectional variable differences between HLCs, HepG2 monolayers and HepG2 spheroids

	HLCs	**HepG2 monolayers**	**HepG2 spheroids**
Positive difference (higher in PHH)	1 020	1 327	682
Negative difference (lower in PHH)	961	956	180
	HLCs versus HepG2 monolayers	**HLCs versus HepG2 spheroids**	**HepG2 monolayers versus spheroids**
Same positive difference	751	500	653
Same negative difference	282	66	142
Multidirectional variable differences	62	11	0

Based on data extracted from volcano plots, positive differences meaning higher expression in PHHs, was most prevalent in HepG2 monolayers and the lowest in HepG2 spheroids (Table 5.9). However, negative differences were higher in HLCs with the lowest in HepG2 spheroids again.

HLCs versus HepG2 monolayers and spheroids had a greater overlap in proteins with a positive compared to negative differences. This suggests that where protein expression is reduced compared to PHHs, it is more common among hepatic phenotypes but when proteins have an increased expression greater than PHH this is more specific to each hepatic model.

Comparing proteins with a positive difference in HLCs to a negative difference in HepG2 spheroids and vice versa highlighted 62 proteins of which 19 were higher in HepG2 cells and 43 higher in HLCs. Of the proteins higher in HepG2 cells, many were associated with DNA replication, transcription and translation (Table 5.10). Cell proliferation does not normally occur in PHHs or HLCs seeded at confluence for maturation. Furthermore, HepG2 spheroids have reduced proliferation where the doubling time is increased and cell division is generally limited to cells at the periphery. As HepG2 cells are immortalized, the ability to propagate indefinitely is considered a valuable feature. However, these proteins suggest that continual replication restrict the mimicry of PHHs and mature iPSC-HLCs which have a limited requirement for replication in order to demonstrate a hepatic phenotype.

Table 5.10: Multidirectional variable differences between HLCs versus HepG2 monolayers (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement	Fold change
Proteins higher in HepG2 monolayers compared to HLCs			
60S ribosomal protein L13	P26373	Ribonuclear protein involved in translation	2.835
Midasin	Q9NU22	Nuclear chaperone for maturation and nuclear export of pre-60S ribosome subunits	2.734
Zinc phosphodiesterase ELAC protein 2	Q9BQ52	Mitochondrial tRNA 3'-processing endonuclease activity (tRNA maturation)	2.052
Pericentriolar material 1 protein	Q15154	Centrosome assembly and function Correct localization of centrosomal proteins	2.443
Integrin-linked kinase-associated serine/threonine phosphatase 2C	Q9H0C8	Regulation of cell cycle progression via dephosphorylation	2.513
Condensin complex subunit 1	Q15021	Regulatory subunit of condensin complex required for conversion of interphase chromatin into mitotic-like condense chromosomes	3.966
Proteins higher in HLCs compared to HepG2 monolayers			
Glyceraldehyde-3-phosphate dehydrogenase	P04406	Role in glycolysis and nuclear events (transcription, RNA transport, DNA replication and apoptosis)	2.716
Arylamine N-acetyltransferase 1	P18440	Participates in the detoxification of a plethora of hydrazine and arylamine drugs	4.133
TBC1 domain family member 2A	Q9BYX2	Acts as GTPase-activating protein for RAB7A (inhibition of cadherin degradation and reduced cell-cell adhesion)	3.360
Protein NDRG1	Q92597	Stress-responsive protein involved in hormone responses, cell growth and differentiation	8.408
Galactocerebrosidase	P54803	Hydrolyzes the galactose ester bonds Enzyme responsible for lysosomal catabolism	38.647
cAMP-dependent protein kinase inhibitor	Q9C010	Competitive inhibitor of cAMP-dependent protein kinase (negative regulation of protein kinase activity)	8.781

Proteins increased in HLCs were found to be involved in catabolism or synthesis and metabolic pathways. Proteins included gamma-glutamyltranspeptidase 1 (P19440) which is a a part of cellular antioxidant defences, kynurenine-oxoglutarate transaminase 3 (Q6YP21), serine racemase (Q9GZT4), aminopeptidase B (Q9H4A4), probable phospholipid-transporting ATPase IF (Q9Y2G3), protein-glutamate O-methyltransferase (Q9H993), phosphoglycerate mutase 1 (P18669), dihydroxyacetone phosphate acyltransferase (O15228), N-acetylgalactosamine-6-sulfatase (P34059), gamma-glutamyltranspeptidase 1 (P19440), receptor-type tyrosine-protein phosphatase (Q12913), di-N-acetylchitobiase (Q01459) and polypeptide N-acetylgalactosaminyltransferase 3 (Q14435). These proteins, utilized in cellular defence mechanisms, could further implicate the time point of collection for proteomics as a contentious one for HLCs.

When comparing proteins with an enhanced expression in HLCs to suppression in HepG2 spheroids and vice versa only 11 variably expressed proteins were highlighted. Two of which were higher in HepG2 spheroids and 9 in HLCs. The reduced number of proteins differentially expressed compared to HepG2 monolayers is likely associated with the fact that HepG2 spheroids have significantly less proteins differing from PHH (862) which reduces the protein cohort for comparison. Of the variably expressed proteins which were increased in HLCs, only two were involved in molecular functions of biological process of noteworthy consequence.

Cathepsin L1 (P07711) is a lysosomal cysteine protease associated with degradation of proteins in lysosomes (TheUniProtConsortium, 2014). Protein degradation and recycling is a normal function within cells. However, increased expression, suggesting increased intracellular protein catabolism, can be linked to cellular stress. The two-fold increase over PHHs and even greater increase over HepG2 cells could reiterate that HLCs collected at Day 35 were undergoing protein degradation or cell stress. The second protein was, deoxyribonuclease-2-alpha (O00115), which plays a role in degradation of nuclear DNA in apoptosis during proper foetal development and is required for definitive erythropoiesis in the foetal liver (TheUniProtConsortium, 2014). A more than two-fold greater expression in HLCs than PHHs suggests a foetal hepatic phenotype of HLCs.

Comparing proteins with enhanced expression in HepG2 monolayers to suppression in HepG2 spheroids and vice versa highlighted no proteins which were variably expressed. As spheroid cultures should merely promote maturation, it makes sense that 95.7% (653 of 682) of proteins up-regulated and 78.9% (142 of 180) of down-regulated proteins in the HepG2 cells variable cultured were the same. Comparing differentially regulated proteins between HLCs, HepG2 monolayers and HepG2 spheroids was only capable of identifying a total of 73 proteins with multidirectional variability.

Based on the total number of proteins, this is only 1.39% (73 of 5231) of the proteins used for analysis. Some of these proteins have provided insight into the specific model. It is important to note that these assessments only compared proteins which were significantly differently express compare to PHHs in the first instance. This does not negate the fact that more proteins could still be differentially regulated between the other model systems. However, here the important feature was the relative relationship to PHHs. Overall this data suggests that, despite varying different degrees of expression compared to PHH, under most circumstances proteins generally vary similarly and to a comparable degree.

5.3 Conclusion

The degree to which the proteome of HLCs, HepG2 monolayers or HepG2 spheroids mimicked that of PHHs was assessed globally with hierarchical clustering, principal component analysis, volcano plots and protein clustering. Collectively this data reveals that none of HLCs, HepG2 monolayers or HepG2 spheroids greatly outperform one another in mimicking PHHs. HLCs appeared to cluster with and correlate closer to PHH than the other two models did to PHHs despite shortcomings in liver specific functions. HepG2 cells appear to be the most limited in hepatic phenotype compared to PHHs. However, there does appear to be an ability to obtain a phenotype closer to PHHs when cultured in 3D as spheroids. This could be due to obtaining morphological characteristics closer to PHHs, enhancing maturation of the liver phenotype in spheroids or a combination of features from both. When comparing the proteomes in terms of ADME proteins, down regulation of the major hepatic metabolizing enzymes in HepG2 cells and HLCs suggest that HepG2 cells and HLCs show greater similarity to one another than to isolated PHHs.

The *in vivo* phenotype of PHHs can be diminished during isolation and qualification procedures after liver biopsy and should also be considered as an influential feature. Furthermore, HLCs demonstrate batch dependent maturation and immortalized cell lines suffer notable phenotypic drift over time. Cellular homeostasis, growth and hepatic maturity exist on a continuum which makes it difficult to repeatedly capture the same point of proteome expression. Therefore, while the conclusions and implications generated here are valid they merely represent a "*snapshot*" of a fluctuating proteome.

What is evident is that by using labelled proteomics with careful preparation steps, over 6 000 proteins per replicate with an overlap of 5 231 proteins could be confidently identified. This quantitative data provided biological insight into the feasibility of use of HLCs, HepG2 cells and HepG2 spheroids for hepatotoxicity testing models. Evidence highlighting minimal differential regulation of proteins in HLCs, HepG2 cells and HepG2 spheroids suggest that if key features are identified which promote the maintenance of the hepatic phenotype of one of these models they could be relevant to other models.

Chapter 6*: In vitro* Hepatotoxicity

6.1 Materials and Methods

6.1.1 Study Drugs

Many NCEs do not make it to commercialisation as drugs due to not meeting the required levels of safety and efficacy while others are withdrawn after commercialisation due to unsatisfactory adverse effects including drug induced liver injury (DILI). There are numerous mechanisms by which drugs can show undesirable toxic effects: firstly, intrinsic toxicity occurs where the drug compound reacts directly with the cells or receptors and thereby elicits a toxic response or destroy the cell directly through compromising membrane integrity. Secondly, drugs can be metabolised within a cell to form a reactive compound that elicits a toxic effect through the reactivity of the modified compound. This can again be through direct binding to molecules or inhibition of critical pathways required for normal survival or function. Thirdly, induction of responses within the cell that alter cellular or tissue homeostasis which then induces an immune response that could include inappropriate inflammatory processes.

DILI is a major concern for new drug classes under development due to the unknown mechanisms of action and lack of safety data and metabolic product knowledge. For this study, drugs were selected from therapeutic classes based on their potential for DILI concern as defined by the Food and Drug Administration (FDA). Three therapeutic classes in the top 10 of highest DILI concern, which took clinical practice and acceptability of hepatotoxicity profile into account, were selected. The selected drugs were either withdraw from the market or used with restrictions. For the purposes of validation, comparison to a drug with no-DILI concern was also selected.

Figure 6.1 gives a graphical representation of different classes of therapeutic drugs with a grading of the degree of concern that each drug class could induce DILI. As could be expected, drugs that have cytotoxic activity such as antineoplastic drugs show the highest concern. However, NSAIDs, antivirals, antidepressants and antibacterial drugs that are all very commonly prescribed also show a high potential for adverse effects including DILI. Table 6.1 summarised the drugs chosen for this study with the year of withdrawal and the FDA rating of concern for inducing DILI. It should be noted that the concern is for drugs administered orally and two of these drugs are still currently available as topical formulations.

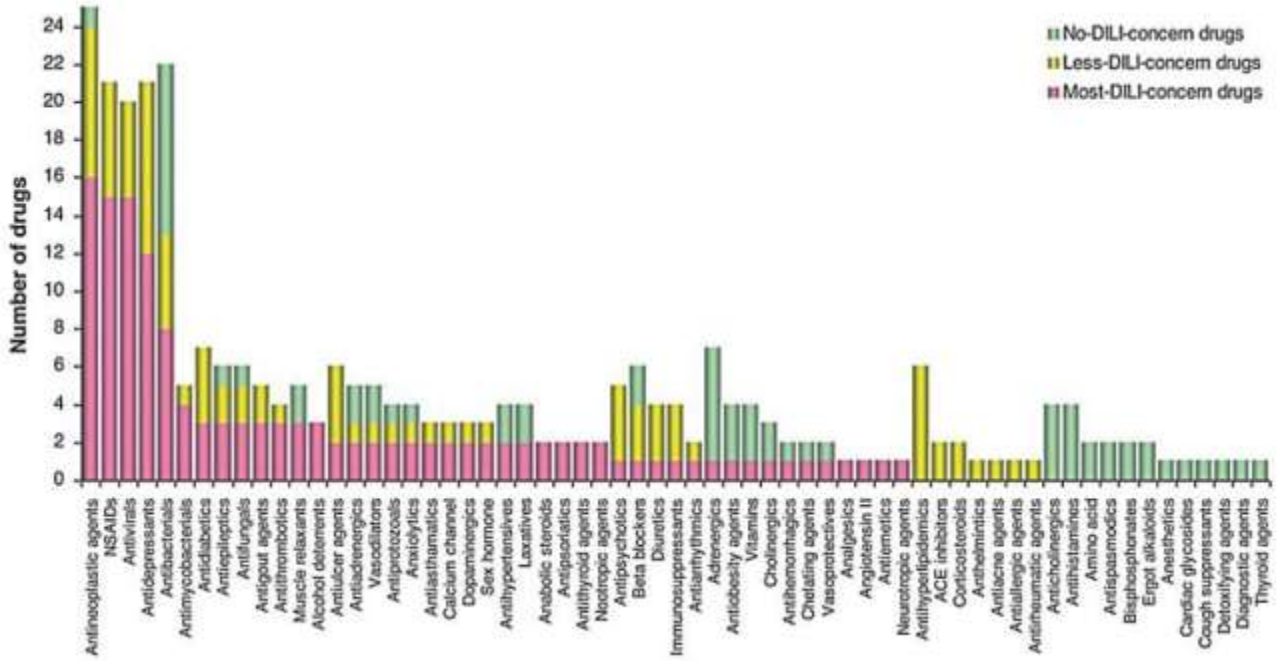

Figure 6.1: Graphical representation of therapeutic classes subdivided into drugs of most, less and no drug-induced liver injury (DILI) concern. (Reprinted with permission from Elsevier, license number 3961470716432, (Chen M, Vijay V et al., 2011))

Table 6.1: Study drugs with their associated therapeutic class, FDA listing and market status

Drug	Therapeutic class	FDA listing	Market status
Bromfenac	NSAID	Most DILI concern (2nd)	Withdrawn (1998)
Ketoconazole	Antifungal	Most DILI concern (9th)	Withdrawn (2013)
Troglitazone	Antidiabetic	Most DILI concern (7th)	Withdrawn (2000)
Diphenhydramine	Antihistamine	No DILI concern	On the market

**Drugs were withdrawn for oral administration, Bromfenac has been remarketed for ophthalmic applications and various topical preparations of Ketoconazole are available*

6.1.2 Preliminary Cytotoxicity

Immortalized hepatocytes exhibit discontinuous phenotypic characteristics, functional properties and genetic instability over time (Duret C, Gerbal-Chaloin S *et al.*, 2007). However, ease of culture and the availability of extensive comparative literature means that HepG2 monolayers remain in use for hepatotoxicity testing. Determining whether known *in vivo* hepatotoxins could exhibit predictable hepatotoxicity to HepG2 cultures was considered essential prior to global proteomic analysis.

Cells were enumerated using a modified sulforhodamine B (SRB) assay. SRB, an anionic aminoxanthene dye, stoichiometrically binds to basic amino acid residues to quantitatively determine protein content which is used as an inference for cell numbers or cell viability (Vichai V and Kirtikara K, 2006). HepG2 cells (Cellonex, South Africa; CHG2-C) were seeded at 2 x 10^4 cells per well (100 µl) of a 96-well plate in EMEM with 10% FBS and left to adhere overnight. Cells were exposed, in triplicate, to half-log concentrations (0.1 - 100 µM; 100 µl) of each drug (50 mM stock solutions in DMSO) made up in EMEM. Vehicle-treated controls (0.1% DMSO) were included to account for the effects of the organic solvent on cells growth. Following 72 hours incubation at 37°C, 50 µl of 50% (w/v) trichloroacetic acid (TCA) was added to each well to fix cells and left overnight at 4°C. Media was removed, plates washed with running water and left to dry at 60°C. One hundred microliters of SRB solution (0.057% (w/v) in 1% (v/v) acetic acid) was added to each well and plates incubated at room temperature for 30 minutes. Plate were washed four times with 1% (v/v) acetic acid, left to dry at 40°C and solubilised with 200 µl Tris-base solution (10 mM; pH 10.5) on an orbital gyratory shaker. Absorbance was quantified at a dual wavelength of 510 nm and 630 nm using an Elx800 Bio-Tek microplate reader.

6.1.3 Multi-parametric Assay Model

To further elucidate hepatotoxic potential, a modified version of an *in vitro* mechanistic toxicity screening model using cultured hepatocytes was conducted (Van Tonder JJ, 2011). Cultured HepG2 cells (2 x 10^4 cells per well), were exposed to five half-log drug concentrations (1 - 100 µM; 100 µl), in duplicate, for 24 and 72 hours. Parameters evaluated, on a single 96-well plate (Figure 6.2), were cell enumeration (CV), reactive oxygen species (ROS), intracellular reduced glutathione (GSH), mitochondrial membrane potential (MMP), intracellular fatty acid accumulation (FAA or steatosis) and apoptosis.

For ROS, GSH, MMP and steatosis assays, medium was aspirated and cells incubated with 100 µl of dye (Table 6.2) in Hank's Balances Salt Solution (HBSS) for 2 hours at 37°C. Plates were read, at their respective excitation and emission wavelengths, using a FLUOstar Optima micro-plate reader. Twenty-five microliters of 50% (w/v) TCA was added to all non-ratiometric assays.

For caspase-3/7 activity, cells were lysed with 25 µl of lysis buffer (10 mM HEPES, 5 mM CHAPS, 2 mM EDTA, with 1 mM PMSF and 5 mM β-mercaptoethanol), on ice for 40 minutes. Following lysis, 100 µl of assay buffer (20 mM HEPES, 2 mM EDTA, with 5 mM β-mercaptoethanol and 5 µM Ac-DEVD-AMC substrate) was added and plates incubated at 37°C for 4 hours. Fluorescence was read and plates incubated at 4°C overnight for continuation of the SRB assay for cell enumeration and fluorescence normalisation.

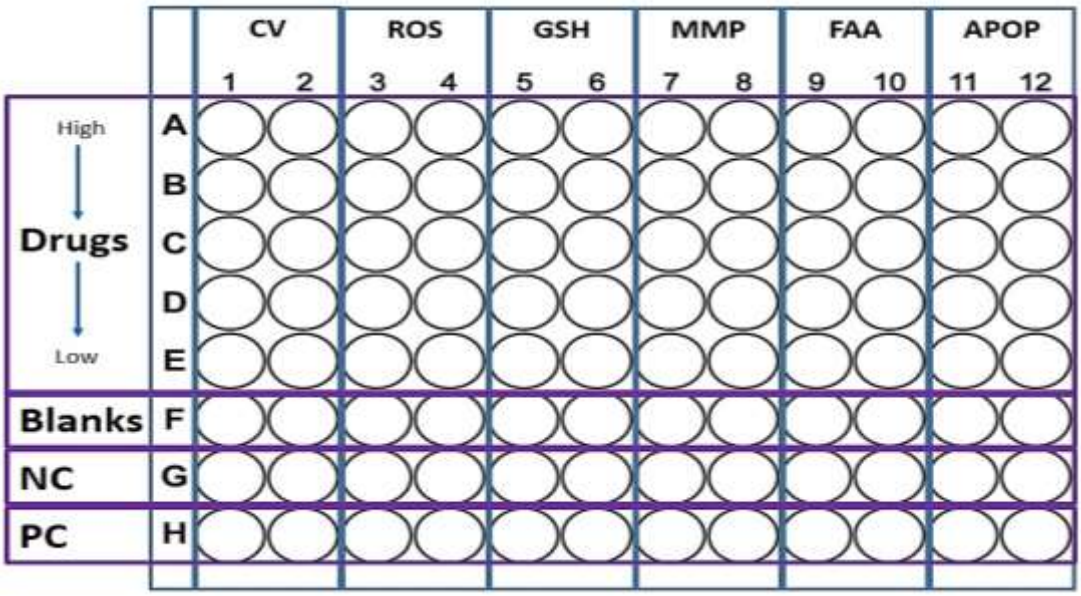

Figure 6.2: Plate layout for the multi-parametric 96-well plate model. CV: cell enumeration, ROS: reactive oxygen species generation, GSH: reduced intracellular glutathione, MMP: mitochondrial membrane potential, FAA: intracellular fatty acid accumulation, APOP: apoptosis measured via caspase-3/7 activity, NC: negative control, PC: positive control.

Table 6.2: Parameters measured using the multi-parametric assay model showing the dye, dye concentration, excitation-emmission wavelengths and positive controls

Parameter	Dye	Concentration	Excitation-emission (nm)	Positive control
CV	SRB	0.057% (m/v)	Absorbance	Rotenone (100 nM)
ROS	H_2-DCF-DA	10 µM	485 / 520	p-Chloranil (5 µM)
GSH	Monochlorobimane	16 µM	355 / 460	N-ethylmalemide (50 µM)
MMP	JC-1	10 µM	485 / 520 (green) 492 / 590 (red)	Rotenone (100 nM)
Steatosis	Nile red	1 µM	544 / 590	Oleic acid (200 µM)
Apoptosis	Ac-DEVD-AMC	4 µM	355 / 460	Rotenone (100 nM)

JC-1 is a ratiometric dye with green fluorescence representing JC-1 monomers in de-energized mitochondria, while red fluorescence represents JC-1 aggregates in energized mitochondria with a higher membrane potential

6.1.4 Statistics for Plate-based Assays

Assays, in 96-well plates, to determine the IC_{50} and hepatotoxic potential, were performed at least in triplicate. Each repeat included internal duplicates or triplicates for blanks, untreated, positive, vehicle-treated controls and study drugs (minimum of n = 6 or n = 9). Results were obtained as fluorescence (100 - 50 000) and/or absorbance (0 - 5.0) units which were blank-adjusted (except for the ratiometric MMP assay) and expressed as a percentage or ratio of untreated controls.

Data was expressed as dose-response curves (nonlinear regression, normalized-variable slope) to determine the IC_{50} or as column graphs for the hepatotoxicity model, using GraphPad Prism 5.0. Due to the low number of internal replicates, Kruskal-Wallis or Mann-Whitney statistical tests were performed. Results, where appropriate, were expressed as a mean ± standard error of the mean (SEM) and a p-values of less than 0.05 was considered as indicator of significance.

6.2 Results and Discussion

6.2.1 *In vitro* Hepatotoxicity

In an attempt to reduce FBS concentration in the culture medium, the influence of FBS concentration on cell density was initially assessed using morphology and cell enumeration by the SRB assay. Ten percent FBS conferred a slight advantage with notably more consistent morphology and phenotype in passaging (Figure 6.3). Therefore, further experimentation was done using 10% FBS.

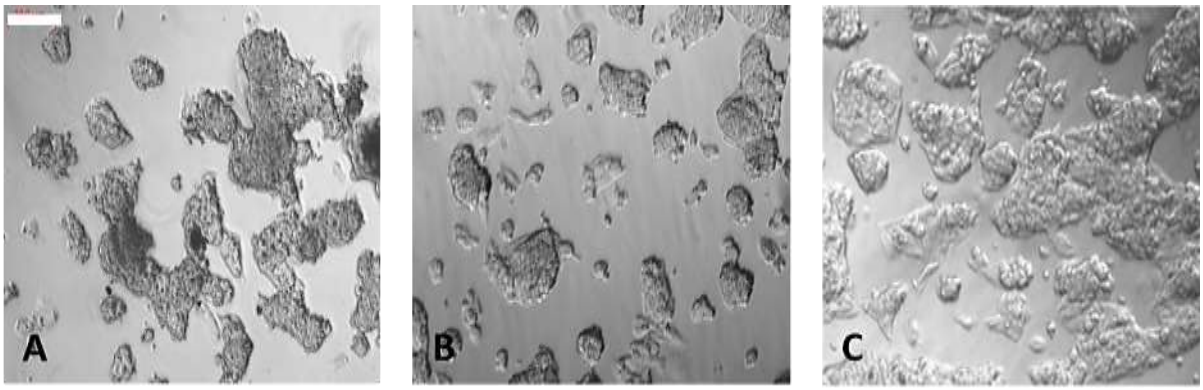

Figure 6.3: Phase contrast images using a Zeiss inverted Axiovert CFL40 microscope of HepG2 cell monolayers cultured in EMEM supplemented with A) No FBS illustrating a rounded morphology indicative of cellular stress, B) Five percent FBS and C) Ten percent FBS illustrating a higher cellular density. Scale bar: 200 µm.

Accurately defining safety profiles of NCEs and INDs relies on the ability of models to predict toxicity. Reactive, electrophilic compounds derived directly from parent drugs or as a consequence to altered cellular radical production are involved in initiating toxicity. Toxicity is often associated with reversible or irreversible protein adduct formation and downstream damage. The formation of reactive metabolites is dependent on both phase I and phase II biotransformation reactions (Williams DP and Park BK, 2003). The liver is, both biochemically and developmentally, one of the most widely studied tissues. Two irreversible stages of cell differentiation involve organogenesis and determination where expression of liver specific function happens. A third reversible stage is modulation which occurs during regeneration and transitions from foetal to adult phenotypes (Kelly JH and Darlington GJ, 1989).

Assessing liver phenotype modulation of the HepG2 cell phenotype can be influenced by cell culture density. Low density cultures (2×10^5 cells/cm^2) exhibit a rapid doubling time and a foetal phenotype. Whereas high density cultures (1×10^6 cells/cm^2) have a prolonged doubling time with increased levels of adult isoenzymes. HepG2 cells, depending on culture duration, can exhibit features characteristic of foetal through to adult or from stationary to regenerative transitions which can be advantageous or limiting for *in vitro* studies (Kelly JH and Darlington GJ, 1989).

Ketoconazole and Troglitazone, which are prototypical hepatotoxic drugs, demonstrated an IC_{50} in the high micromolar range (Table 6.3 and Figure 6.4). However, Bromfenac, which is highly hepatotoxic *in vivo*, mimicked the outcomes observed for non-hepatotoxic Diphenhydramine with the inability to demonstrate an IC_{50} when exposed to concentrations of up to 100 µM.

HepG2 cells have been used for long term studies requiring hepatic cells due to their ease of culture and application for cytotoxicity testing while apparently retaining some properties of primary hepatocyte phenotypes. Reduced transcriptome levels of basal expression of phase I and phase II biotransformation enzymes in HepG2 cells compared to PHHs (LeCluyse EL, Witek RP *et al.*, 2012) play a critical role in the accurate assessment of hepatotoxic drugs. Notable divergent aspect of the HepG2 phenotype include lower expression profiles of various essential phase I enzymes such as CYP2D6 with a relative abundance of 5%. Contrasting the low CYP2D6 abundance is the overexpression of metabolic enzymes such as CYP2W1 in HepG2 cells (Guo L, Dial S *et al.*, 2010).

Phenotypic shifts associated with different passages and cell stocks were noted. This influenced the rate of passage, ability to produce spheroids, abundance and expression of enzymes as well as general intracellular functions. These factors highlight the inadequacies of long-term cultured, immortalized cell lines in cytotoxicity predictions.

Table 6.3: IC_{50} values for study drugs in HepG2 cell monolayers

Drug	Therapeutic class	IC_{50} (µM)
Bromfenac	NSAID	Not converged
Ketoconazole	Antifungal	33.11 ± 2.881
Troglitazone	Antidiabetic	23.93 ± 1.086
Diphenhydramine	Antihistamine	Not converged

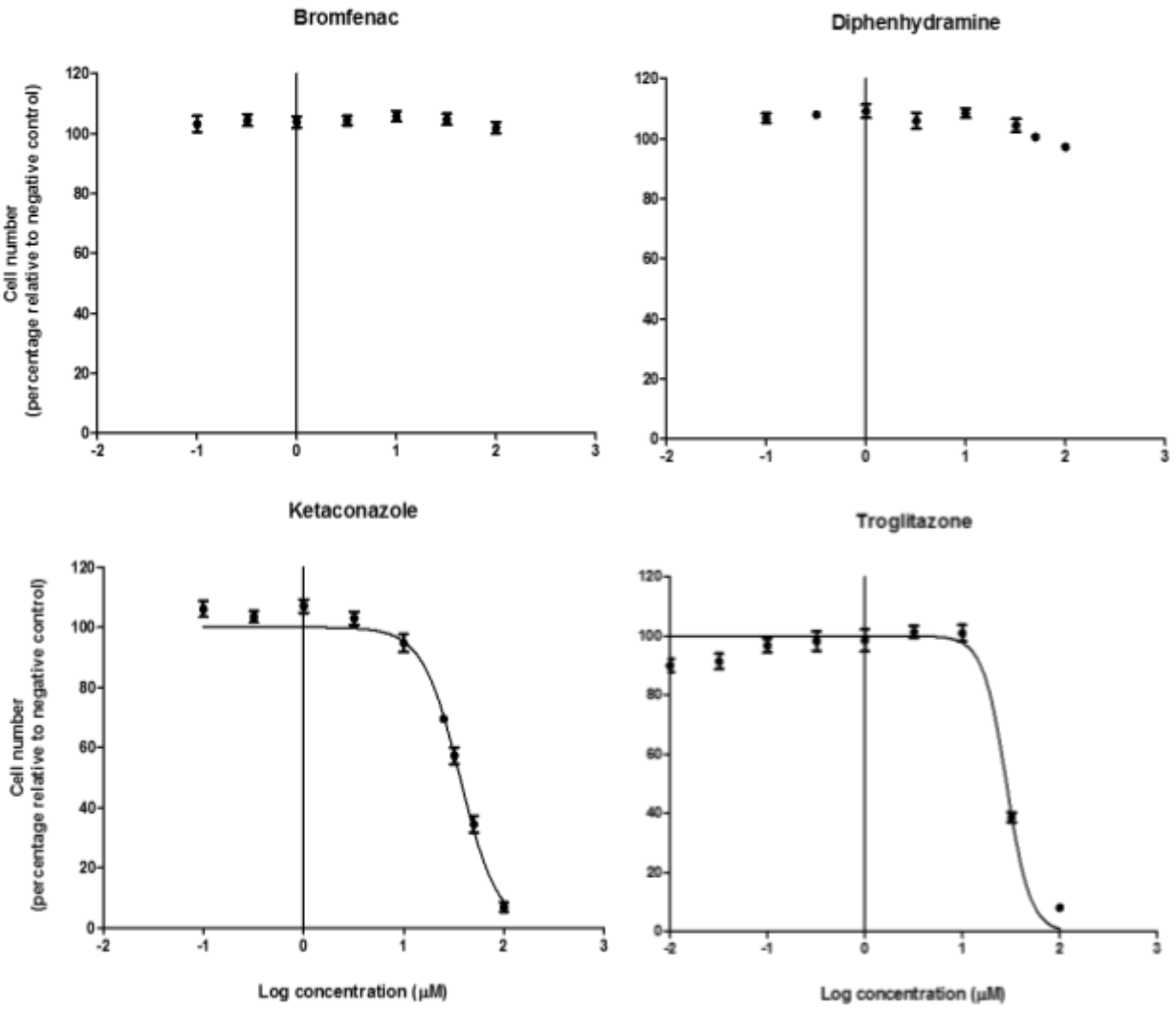

Figure 6.4: Dose-response curves generated using the SRB assay for Bromfenac, Diphenhydramine, Ketoconazole and Troglitazone.

The interplay between factors such as ROS, MMP, GSH, steatosis and executioner caspase activity all contribute to disrupted cellular homeostasis. These can be critical initiating events in cell death with apoptosis being believed to be the major form of cell death associated with hepatic injury. Oxidative damage can be via direct drug action or induction of ROS. Targets of oxidative damage include intracellular biomolecules such as DNA, lipids and proteins which alters both their structure and function and depletes GSH creating a deficiency in defence mechanisms.

Mitochondrial-mediated toxicity can disrupt intracellular calcium, generate ROS, and induce protease and endonuclease activation and cellular dysfunction. Additionally, drugs can directly induce down regulation of anti-apoptotic gene products, enable mitochondrial permeability transition (MPT) and promote DNA fragmentation promoting the apoptotic cascade (Feldmann G, 1997; Gómez-Lechón MJ, Tolosa L *et al.*, 2010; Jones DP, Lemasters JJ *et al.*, 2010).

The multi-dimensional and interactive nature of hepatic injury requires the assessment of all of these factors, which were assessed simultaneously. Considering the time-dependant nature of cellular responses, HepG2 cells were exposed to five half-log concentrations of each drug for 24 and 72 hours. Despite being implications in severe hepatotoxicity, HepG2 cell monolayers were unable to elucidate features associated with Bromfenac-induced hepatotoxicity (Figure 6.5).

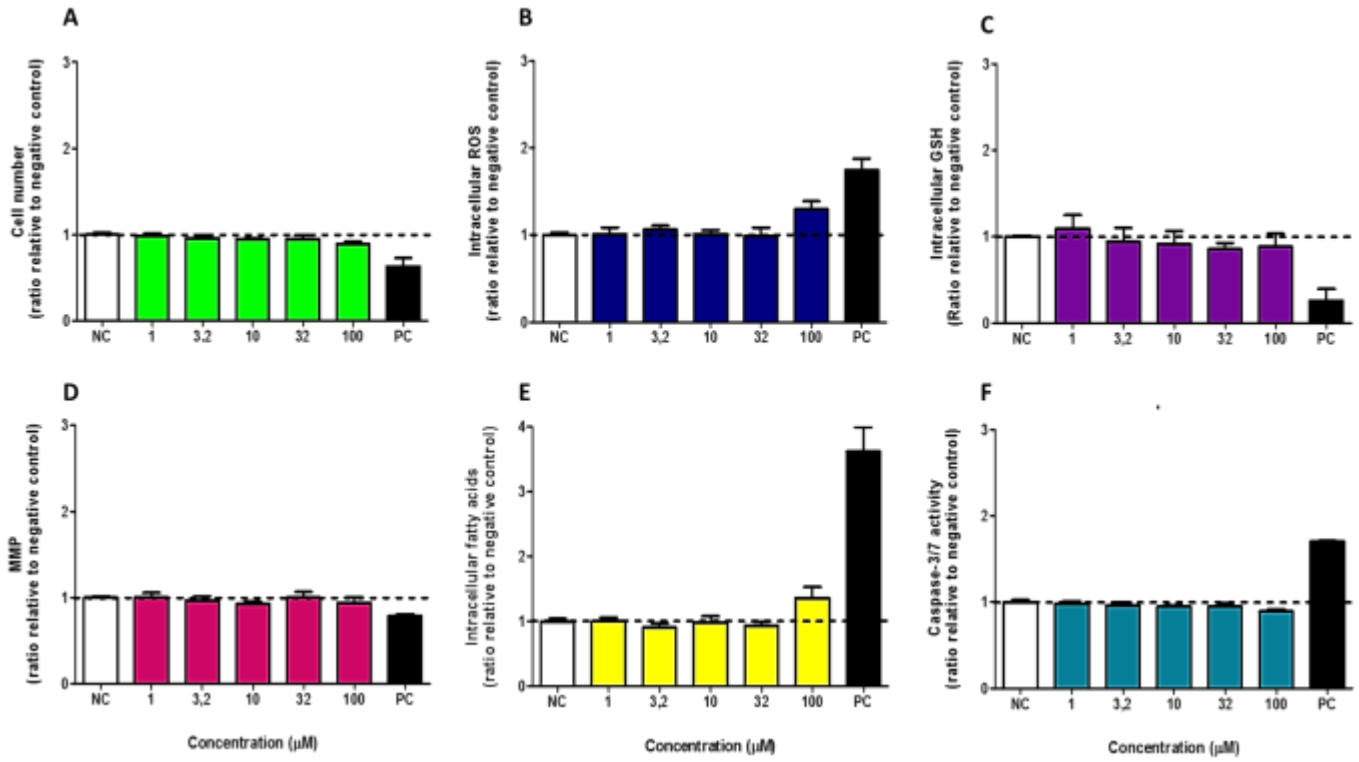

Figure 6.5: Bar graphs for HepG2 cells exposed to Bromfenac of A) Cell enumeration, B) Intracellular ROS, C) Intracellular GSH, D) MMP, E) Steatosis and F) Caspase-3/7 activity. NC: negative control, PC: positive control. Statistically, each positive control was significantly reduced or increased as required for each assay. Significance (*) was defined as $p < 0.05$.

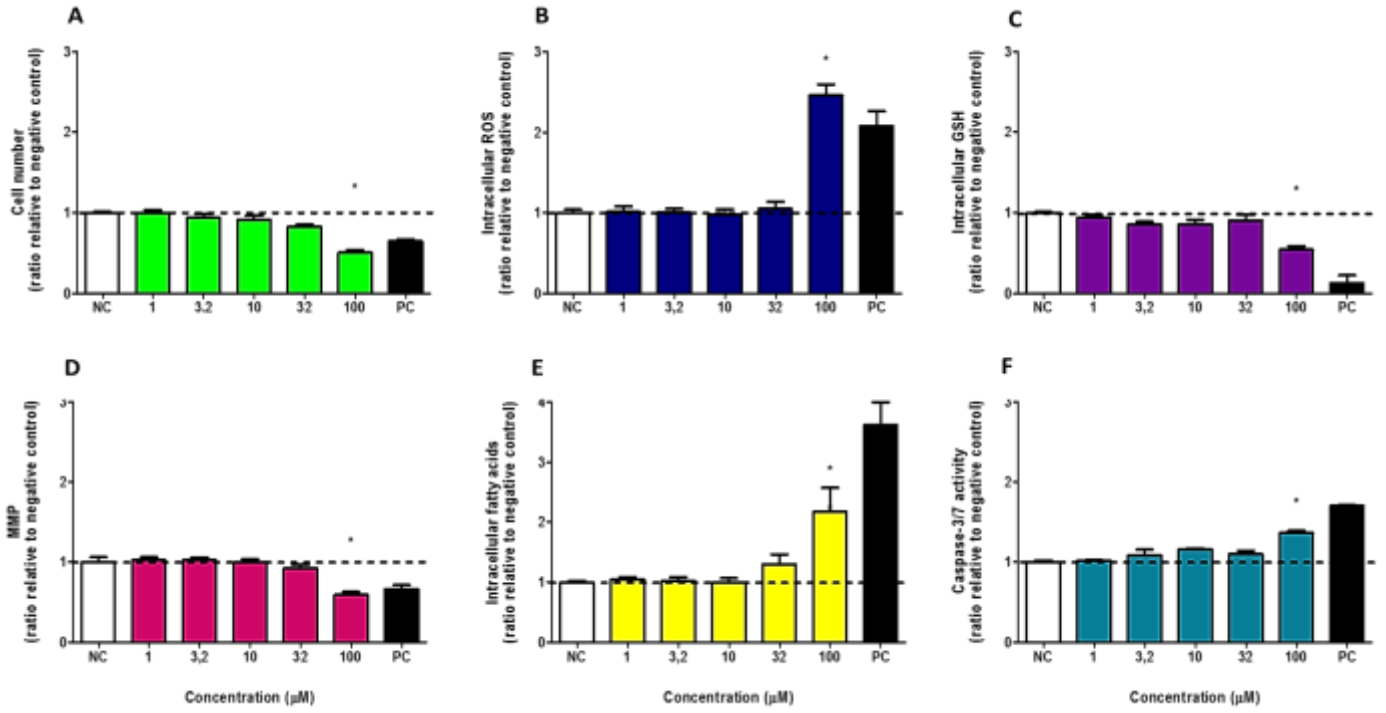

Figure 6.6: Bar graphs for HepG2 cells exposed to Ketoconazole of A) Cell enumeration, B) Intracellular ROS, C) Intracellular GSH, D) MMP, E) Steatosis and F) Caspase-3/7 activity. NC: negative control, PC: positive control. Statistically, each positive control was significantly reduced or increased as required for each assay. Significance (*) was defined as $p < 0.05$.

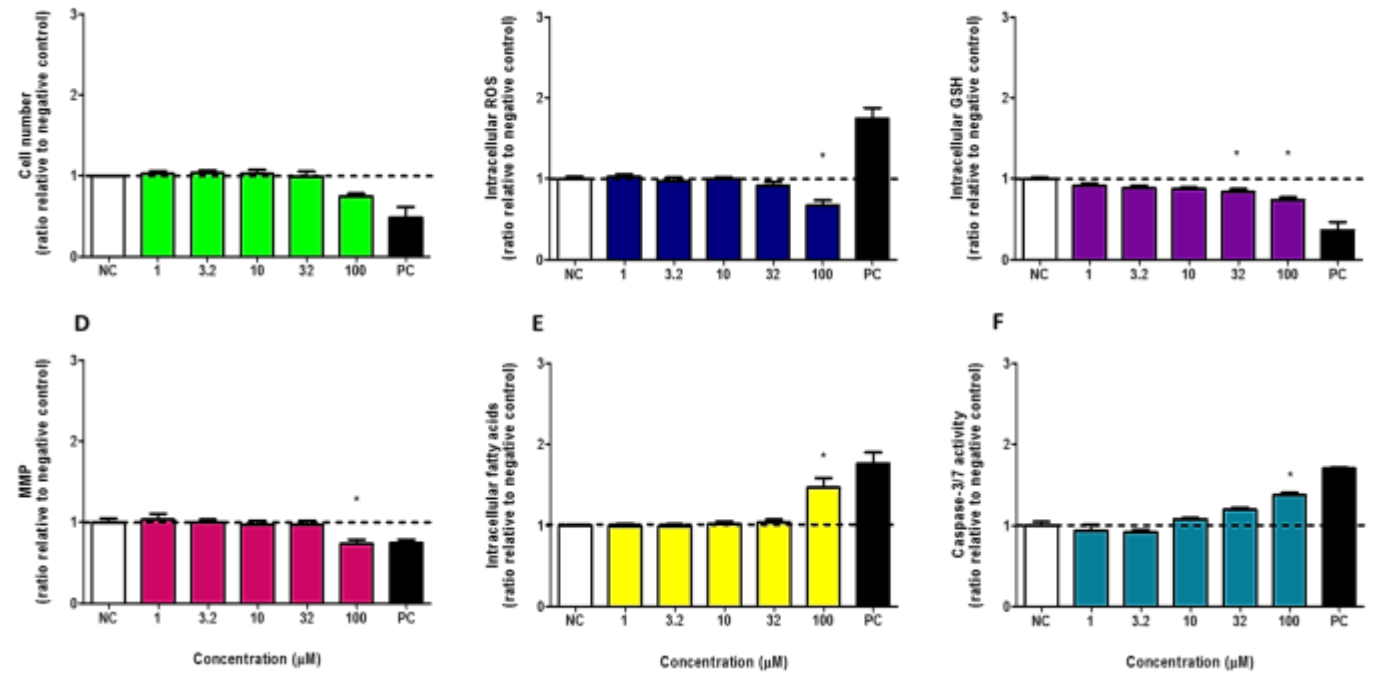

Figure 6.7: Bar graphs for HepG2 cells exposed to Troglitazone of A) Cell enumeration, B) Intracellular ROS, C) Intracellular GSH,D) MMP, E) Steatosis and F) Caspase-3/7 activity. NC: negative control, PC: positive control. Statistically, each positive control was significantly reduced or increased as required for each assay. Significance (*) was defined as $p < 0.05$.

Ketoconazole and Troglitazone produced similar dose and time dependent changes. Slight decreases in cell viability and MMP with a consequent increase in caspase-3/7 activity were observed. Intracellular GSH decreased while fatty acid accumulation increased at high drug concentrations (Figure 6.6 and 6.7). The exception to the similar responses was that intracellular ROS was increased in response to Ketoconazole but depleted in response to Troglitazone. Both of these drugs can be directly hepatotoxic but generally induced greater toxicity once metabolised to their corresponding reactive metabolites. Proteomic investigation was conducted prior to detailed discussion of associated pathways and networks affected by drug exposure.

Diphenhydramine, a non-hepatotoxic antihistamine, used for the purposes of validation displayed no toxicity or alterations in any of the assays conducted (Figure 6.8). This is in keeping with the non-hepatotoxic characteristics of Diphenhydramine. However, assay results were similar for Bromfenac which is hepatotoxic *in vivo*. What is apparent is that, while validation of model systems is a necessity, here it appears unable to fulfil its purpose due to the numerous shortcomings of HepG2 monolayer cultures. Alternatively, this could merely illustrate the limitations of endpoint assays in determining hepatotoxicity which may potentially be overcome when assessing the whole proteome.

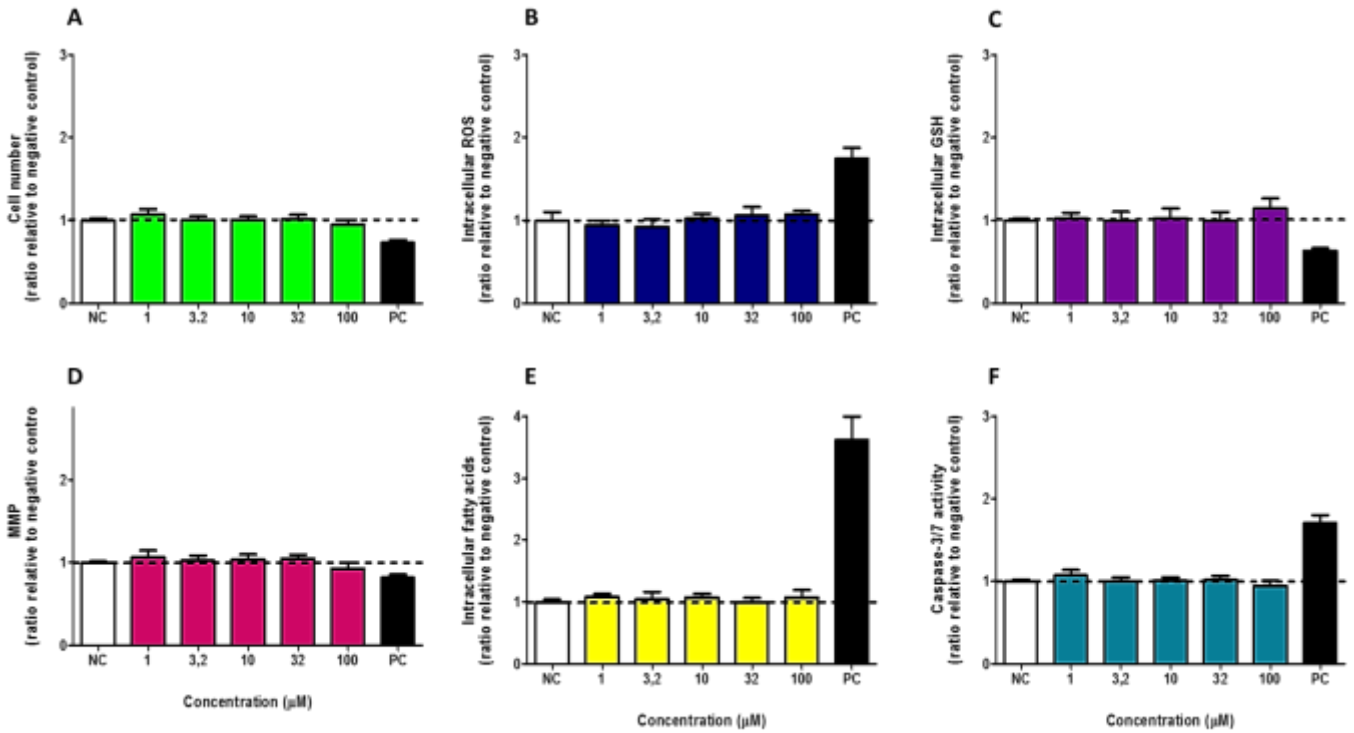

Figure 6.8: Bar graphs for HepG2 cells exposed to Diphenhydramine of A) Cell enumeration, B) intracellular ROS, C) Intracellular GSH, D) MMP, E) Steatosis and F) Caspase-3/7 activity. NC: negative control, PC: positive control. Statistically, each positive control was significantly reduced or increased as required for each assay. Significance (*) was defined as $p < 0.05$.

6.3 Conclusion

Despite HepG2 cells being considered largely metabolically incompetent, they are commonly used in hepatotoxicity screening strategies. *In vitro* mechanistic assays demonstrated various possible cytotoxicity effects, but only at very high concentrations, for Troglitazone and Ketoconazole but not for the well documented hepatotoxin, Bromfenac. *In vitro* hepatotoxicity was required to define sub-toxic drug concentrations (IC_{20}) for differential protein expression analysis using mass spectrometry based proteomics. As no IC_{50} could be determined for Bromfenac and Diphenhydramine the calculated C_{max} following a maximum oral dose of each of these drugs was determined and used for proteomic shift experiments.

These analyses were only conducted in monolayer cultures as the efficacy and permeability of each dye for the different mechanistic studies in 3D models had not been elucidated. Protein content, as an inference for cell numbers, using the BCA assay was done to confirm that using the C_{max} or IC_{20} concentration obtained in 2D assays was applicable in the 3D culture technique prior to collecting samples for gel and MS-based proteomics.

This cytotoxicity and mechanistic data highlighted the inability of HepG2 cells with this model to predict *in vitro* hepatotoxicity of known *in vivo* hepatotoxins. Enhanced predictive power using HepG2 cells might be achieved by culturing cells in 3D or by using a more sensitive endpoint such as global proteomics. Alternatively, HepG2 cells need to be abandoned in favour of other sources of hepatic cells. A consideration which could have been advantageous is to use a toxic/non-toxic pair within a specific therapeutic class instead of an independent non-hepatotoxic drug. However, if global proteomics provides insight into useful application, this could still be conducted.

Chapter 7: Proteomic Profiling of Cytotoxicity

7.1 Materials and Methods

7.1.1 Drug Exposure

The sulforhodamine B assay was used to determine the drug concentrations for which monolayersand spheroids were to be exposed to prior to proteomic profiling. Ketoconazole and Troglitazone, which demonstrated a measureable IC_{50} concentrations using HepG2 cell (Cellonex, South Africa; CHG2-C) monolayers were used at their calculated IC_{20} of 20 µM and 15 µM respectively. As no IC_{50} could be determined for Bromfenac and Diphenhydramine, HepG2 cells were exposed at their calculated C_{max} (70 kg person with 5.5 L blood and 80% oral bioavailability) of 25 µM and 30 µM respectively. HepG2 monolayers were seeded (2.5 x 10^5 cells/well), into 6-well plates and incubated for 4 days. Spheroids (1 x 10^4 cells per 45 µl) were seeded into hanging drops for 4 days for initial aggregation prior to continued culturing in 6-well plates (n = 25 spheroids per well). Monolayers and spheroids were exposed to the drugs from Day 4 and left to incubate at 37°C with 5% CO_2 for 72 hours. Re-exposure was done for an additional 72 hours with protein collection at Day 10.

7.1.2 Gel-based Proteomics

Materials and methods for sample collection, protein preparation and one-dimensional SDS-PAGE were conducted as described in Chapter 4.

7.1.3 Mass Spectrometry-based Proteomics

Materials and methods for preparing proteins from cultured cells including; sample collection and lysis, protein quantitation, protein digestion, peptide TMT labelling (Table 7.1 and 7.2) and SPE extraction clean-up were conducted as described in Chapter 4. RP-HPLC with selected recombination and LC-MS/MS were done using different instrumentation which is detailed below.

Table 7.1: Tandem mass tag labelling to compare HepG2 cells as monolayers and spheroids before and after exposure to Bromfenac and Troglitazone

TMT label	Replicate 1	Replicate 2	Replicate 3
126	Troglitazone spheroids	Troglitazone spheroids	Control monolayers
127	Bromfenac spheroids	Bromfenac spheroids	Bromfenac monolayers
128	Control spheroids	Control spheroids	Troglitazone monolayers
129	Troglitazone monolayers	Troglitazone monolayers	Control spheroids
130	Bromfenac monolayers	Bromfenac monolayers	Bromfenac spheroids
131	Control monolayers	Control monolayers	Troglitazone spheroids

Table 7.2: Tandem mass tag labelling to compare HepG2 cells as monolayers and spheroids before and after exposure to Diphenhydramine and Ketoconazole

TMT label	Replicate 1	Replicate 2	Replicate 3
126	Ketoconazole spheroids	Ketoconazole spheroids	Control monolayers
127	Diphenhydramine spheroids	Diphenhydramine spheroids	Diphenhydramine monolayers
128	Control spheroids	Control spheroids	Ketoconazole monolayers
129	Ketoconazole monolayers	Ketoconazole monolayers	Control spheroids
130	Diphenhydramine monolayers	Diphenhydramine monolayers	Diphenhydramine spheroids
131	Control monolayers	Control monolayers	Ketoconazole spheroids

7.1.4 Reverse Phase HPLC

7.1.4.1 Instrument Parameters and QC

RP-HPLC was done using a Shimadzu HPLC system coupled to a PDA detector. Priming was done for 3 minutes using 50:50 mobile phases A and B (Table 7.3) at a flow rate of 8 ml/minute. An Ascentis C18 HPLC column (150 mm x 4.6 mm, 5 µm; octadecyl bonded phase) was initially washed with 95% mobile phase B for 12 minutes and then with mobile phases A and B (50:50) for 10 minutes at 1 ml/min and pressure stabilized before reducing to 5% mobile phase B.

Table 7.3: Shimadzu HPLC system conditions and settings for RP-HPLC

Binary Solvent Manager	
Mobile phase A (aqueous)	20 mM ammonium formate, pH 10
Mobile phase B (organic)	80% acetonitrile 20 mM ammonium formate, pH 10
Pressure limit: low	0 psi
Pressure limit: high	5000 psi
PDA Detector	
Wavelength range	220 - 400 nm: 220, 254, 280 nm traces recorded
Frequency	4.1667 Hz
Sampling rate	20 points per second
Band width	4, 10 and 20 nm for abovementioned channels
Filter time constant	Normal (0.1 seconds)
Resolution	1.2 nm
Sample Manager	
Autosampler temperature	Room temperature (20°C)
Column temperature	Room temperature (20°C)
Sample loop option	Full loop (100 µl)

Three column washes were conducted using a 12 minute run (Table 7.4). Quality control was done using cytochrome C (1 mg/ml in 10 mM ammonium bicarbonate, 25 µl: Figure 7.1) over the same 12 minute run. Two additional washes were done prior to sample loading.

Table 7.4: Mobile phase gradient for wash steps and cytochrome C standard using RP-HPLC

Time (Minutes)	Flow (ml/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	1.00	95	5
8.00	1.00	40	60
8.10	1.00	5	95
9.50	1.00	5	95
9.60	1.00	95	5

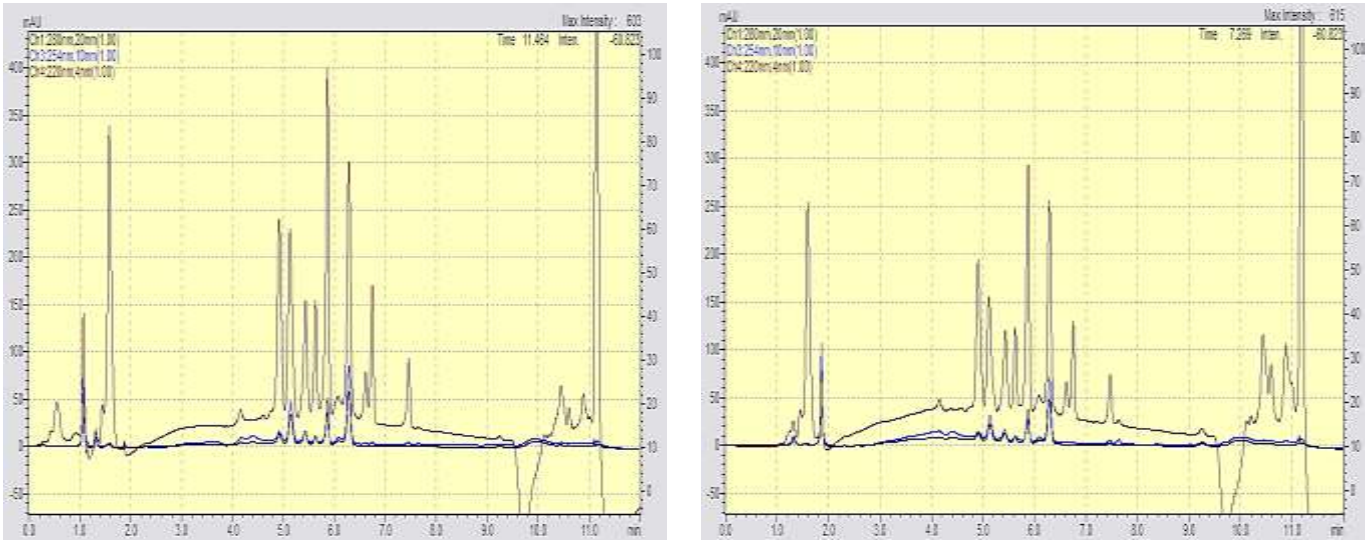

Figure 7.1: Cytochrome C standards run prior to sample fractionation on different days.

7.1.4.2 Sample Peptide Fractionation

Peptide samples, dried down after TMT-labelling were resuspended in 100 µl of 4% acetonitrile with 20 mM ammonium formate (pH 10) and 95 µl injected onto an Ascentis C18 HPLC column. Peptides were eluted at a flow rate of 1 ml/minute using an initial isocratic low organic mobile phase after which the organic phase was increased during a multi-step gradient program over a total 75 minute run (Table 7.5). Chromatography was profiled using three wavelengths of 220, 254, 280 nm with different band widths to enhance sensitivity. Eluting peptide fractions of different times were collected for later recombination in approximately equal peptide concentrations across the gradient (Table 7.6).

Table 7.5: Mobile phase gradient for peptide elution using RP-HPLC for a 75 minute run time

Time (Minutes)	Flow (ml/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	1.00	95	5
5.00	1.00	95	5
63.00	1.00	25	75
65.00	1.00	0	100
70.50	1.00	0	100
70.60	1.00	95	5

Table 7.6: Fraction collection of TMT-labelled peptides separated by RP-HPLC

Time span	Duration of collection (Minutes)	Number of fractions
0 - 15	5	3
16 - 60	1	45
61 - 75	5	3

Fractions were immediately vacuum dried (35°C) and stored at -80°C. To reduce the sample numbers submitted for high resolution mass spectrometry, hydrophobic and hydrophilic fractions were pooled to give 12 fractions based roughly on peptide abundance from 15 to 45 minutes (Figure 7.2). Pooled samples were resuspended in a loading solvent of 2% acetonitrile with 0.1% formic acid and transferred to conical-bottom MS vials for LC-MS/MS analysis.

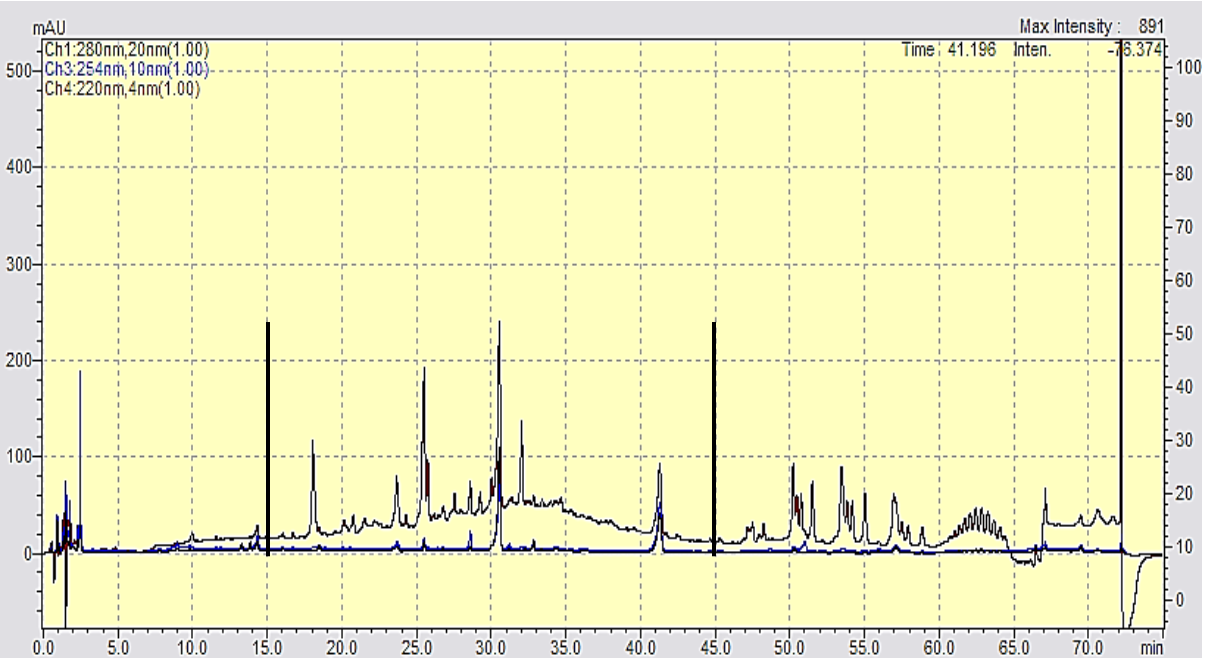

Figure 7.2: Chromatogram of peptide separation for CBT replicate 1, where hydrophilic peptides with affinity for the mobile phase eluted prior to the hydrophobic peptides interacting with the hydrophobic stationary phase. Black lines indicated the time frame of collection of peptide containing fractions.

7.1.5 LC-MS/MS Analysis

Samples were analysed at the Central Analytical Facility, University of Stellenbosch using a Dionex Ultimate 3000 RSLCnano 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 onto a C18 pre-column (Thermo Scientific; 100Å, 5 mm x 300 µm, 5 µm) and a Luna C18 (350 mm x 75 µm, 3.6 µm) analytical column using loading solvent (2% acetonitrile with 0.1% formic acid).

Samples were loaded onto the trap column at a loading pump flow rate of 15 µl/minute from a temperature controlled auto-sampler (7°C) for 5 minutes before eluting onto the analytical column. Peptide separation was performed at 50°C at a rate of 500 nl/minute using mobile phase A (2% acetonitrile with 0.1% formic acid) and mobile phase B (100% acetonitrile). A non-linear gradient (Chromeleon Chromatography Systems 6.0) of 2 - 50% mobile phase B over 105 minutes was used for separation using a 120 minute run time (Table 7.7). LC eluent was introduced into the mass spectrometer via a stainless steel nano-bore emitter.

Table 7.7: Dionex Ultimate 3000 RSLC nanoUPLC system mobile phase gradients

Time (Minutes)	Flow (µl/minute)	Mobile phase A (%)	Mobile phase B (%)
0.00	0.300	98	2
5.00	0.300	95	5
95.00	0.300	65	35
105.00	0.300	50	50
120.00	0.300	98	2

Mass-to-charge (m/z) ratios of positively charged ions, with 2 kV spray voltage, was measured in the Orbitrap mass analyser with data acquired in profile mode. Real-time recalibration by correction of m/z shifts for spectra was done using polysiloxane ions at m/z = 445.12003 (positive) and 371.10024 (positive). MS1 scans were performed using the Orbitrap detector (resolution: 120 000) over the scan range 400 - 1650 m/z with AGC target at 2e5 and maximum Orbitrap injection time of 50 ms.

MS2 acquisition was performed using monoisotopic precursor selection for ions with charge states between 2+ and 6+ with an error tolerance of ± 10 ppm. Selected precursor ions were fragmented by HCD set at a NCE of 32.5% with a stepped NCE of 10% in the quadrupole mass analyser and then excluded from fragmentation once for 30 seconds. Selected fragments were then excluded from fragmentation once for 30 seconds. Fragment ions were detected, using centroid mode (resolution: 60 000) with the AGC target of 1e5 and maximum injection of 120 ms. Raw files collected during the peptide elution were processed with the same bioinformatics pipeline as described in Chapter 4.

7.2 Results and Discussion

It is thought that experimentally curated markers (miR-122, HMGB1 isoforms, keratin-18, glutamate dehydrogenase) which are suggestive of a specific mechanism of toxicity can be used alongside established clinical markers such as ALT/AST, bilirubin and prothrombin time to improve drug safety (Fox D, Morris L *et al.*, 2009). Therefore, this experiment was designed to determine if proteins signatures could be identified that elucidate drug specific toxicity in either 2D or 3D cultures which can be used to in combination with clinical markers.

Protein mass profiles demonstrated no obvious band differences in monolayers exposed to any of the drugs (Figure 7.3). However, spheroids exposed to Ketoconazole and Troglitazone at their respective IC_{20} values demonstrated various band alterations (Figure 7.4) within different molecular mass regions (75 – 100 kDa, 37 - 50 kDa, 25 - 37 kDa). Further investigation into bands was not conducted as whole proteome labelled proteomics could be correlated to the mass ranges where differences were evident.

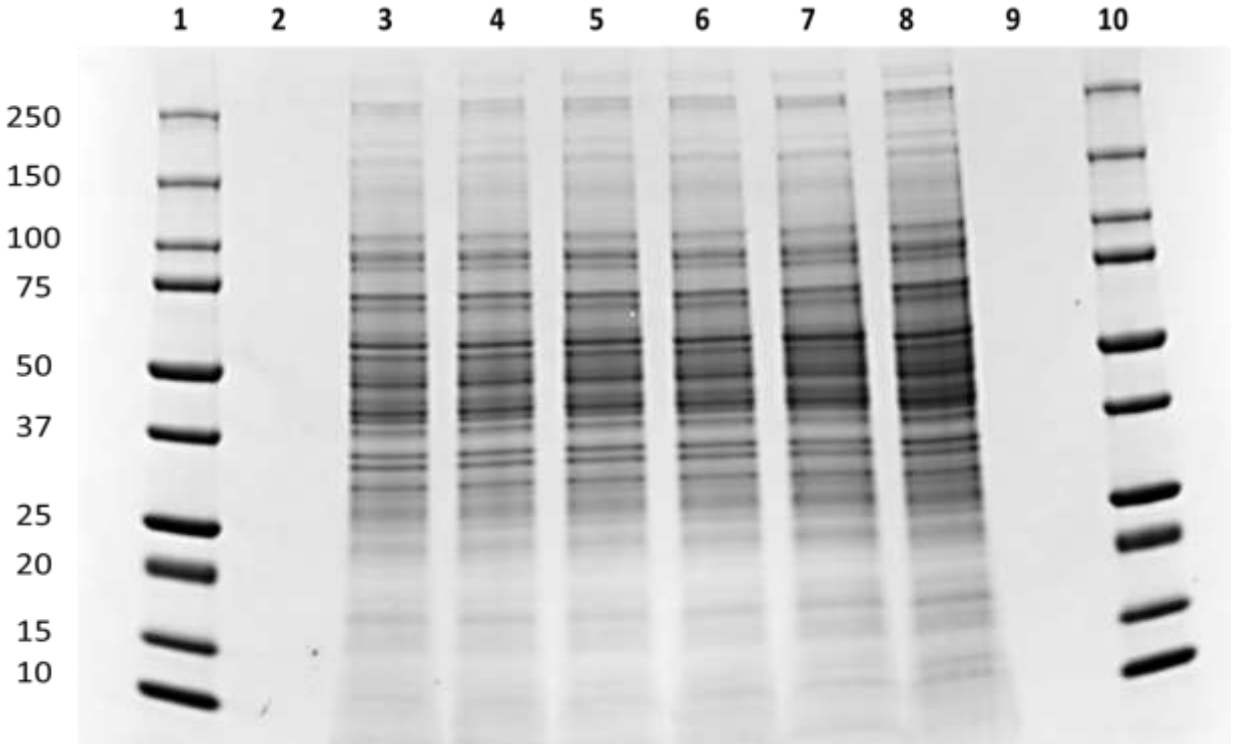

Figure 7.3: Coomassie stained 4-15% Mini-PROTEAN TGX gel of HepG2 monolayers following exposure to selected drugs. 1) Precision Plus Protein Dual Colour standard (10 - 250 kDa), 3 - 4) Control monolayers, 5) Monolayers exposed to Bromfenac, 6) Monolayers exposed to Diphenhydramine, 7) Monolayers exposed to Ketoconazole, 8) Monolayers exposed to Troglitazone and 10) Standard (10 - 250 kDa).

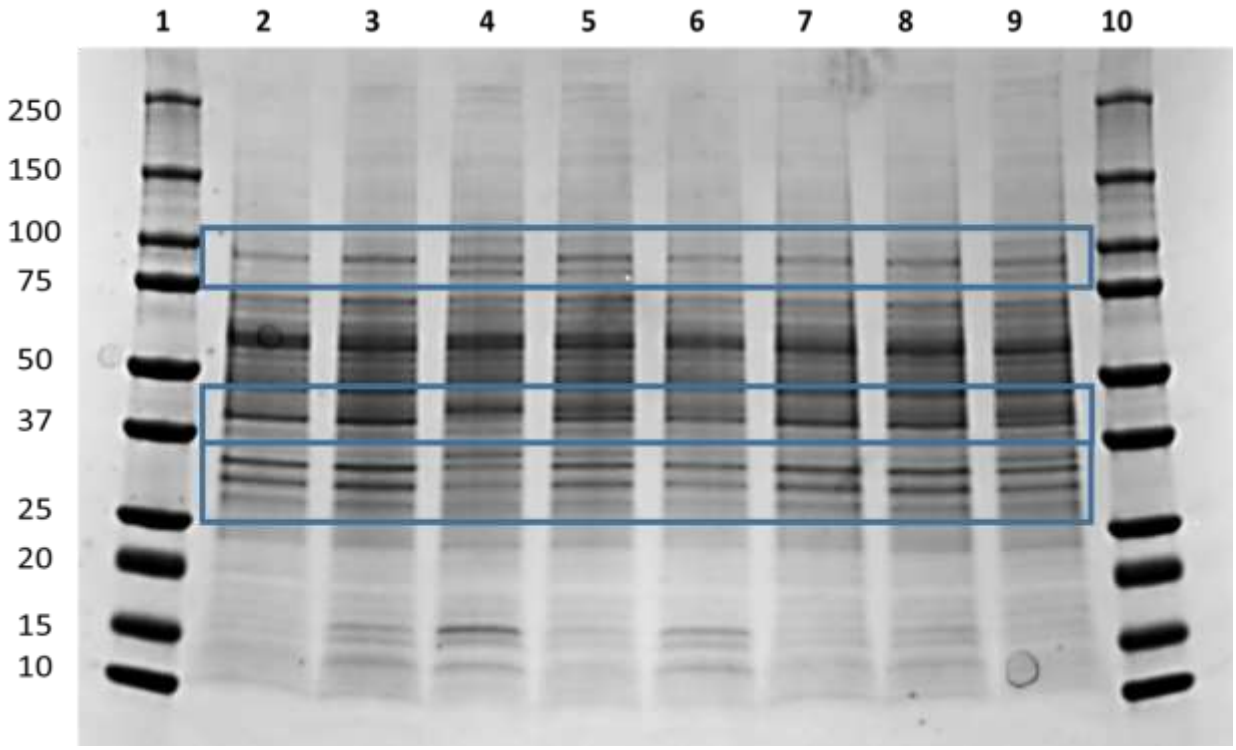

Figure 7.4: Coomassie stained 4-15% Mini-PROTEAN TGX gel of HepG2 spheroids following exposure to selected drugs. Blue boxes highlight band alterations in the 75 - 100, 37 - 50 and 25 - 37 kDa mass ranges. 1) Precision Plus Protein Dual Colour standard (10 - 250 kDa), 2) Control spheroids, 3) Spheroids exposed to Bromfenac, 4) Spheroids exposed to Ketoconazole, 5) Spheroids exposed to Troglitazone, 6) Control spheroids, 7) Spheroids exposed to Bromfenac, 8) Spheroids exposed to Diphenhydramine, 9) Spheroids exposed to Troglitazone and 10) Standard (10 - 250 kDa).

Raw files were inspected as TICs, MS1 and MS2 spectra. As expected, reporter ions of different intensities were evident in the m/z range from 126 - 131 in the MS2 spectra (Figure 7.5). Fraction pooling following HPLC resulted in good orthogonality of separation in each fraction for the LC-MS/MS fractions whereby the total ion intensity was reasonably well distributed across each chromatographic run (Figure 7.6).

Protein identification for CBT replicates 1 to 3 (controls, Bromfenac and Troglitazone labelled in the same 6-plex experiments) was 4 701, 5 030 and 4 932 respectively with an overlap of 3 760 proteins after protein filtering (Table 7.8 and Figure 7.7). Protein identification for CDK replicates 1 to 3 (controls, Diphenhydramine and Ketoconazole labelled in the same 6-plex experiments) was 5 055, 4 971 and 4 680 respectively with an overlap of 3 727 proteins (Table 7.9 and Figure 7.7).

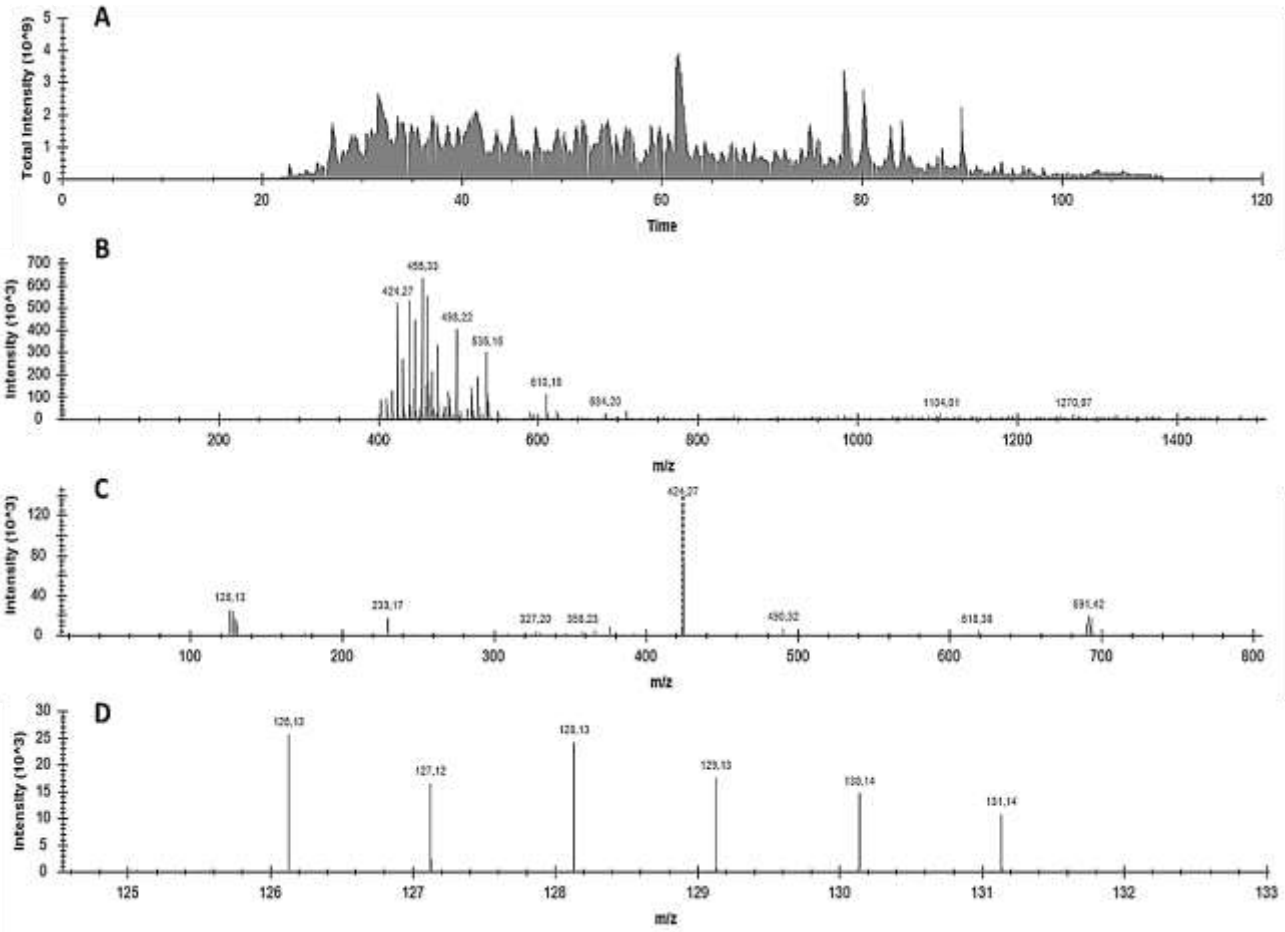

Figure 7.5: LC-MS/MS chromatograms for CBT replicate 1. A) Total ion chromatogram over a 120 minute run, B) MS1 chromatogram, C) MS2 chromatogram, D) MS2 chromatogram zoomed in to the m/z 126 - 131 range illustrating the six reporter ion intensities.

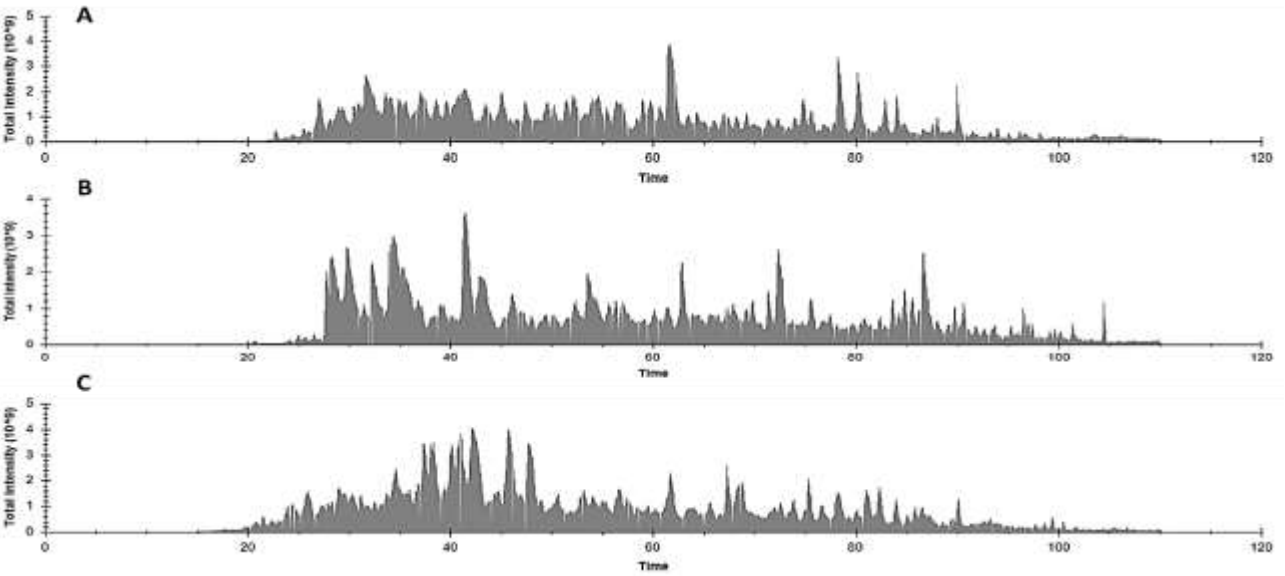

Figure 7.6: Total ion chromatograms for A) Fraction 2, B) Fraction 6 and C) Fraction 9 of CBT replicate 1 demonstrating the variance in total intensity across each 120 minute run.

Table 7.8: Global summary for CBT replicates

	Replicate 1	Replicate 2	Replicate 3
Proteins	4 701	5 030	4 932
Protein group(s): 2+ unique peptides	4 520	4 832	4 748
Protein group(s): 2+ unique peptides, high confidence (100%)	4 361	4 669	4 604
	Protein group(s): 2+ unique peptides, high confidence		
Range of peptide numbers	2 - 258	2 - 280	2 - 273
Range of peptide spectrum matches	2 - 1135	2 - 1676	2 - 1571
Molecular weight range (kDa)	6.24 - 837.78	6.24 - 837.78	6.24 - 837.78
Coverage range	0.17 - 86.22	0.27 - 84.31	0.27 - 84.31
Proteins in replicate 1 to 3	3760		

Table 7.9: Global summary for CDK replicates

	Replicate 1	Replicate 2	Replicate 3
Proteins	5 055	4 971	4 680
Protein group(s): 2+ unique peptides	4 837	4 741	4 495
Protein group(s): 2+ unique peptides, high confidence (100%)	4 665	4 610	4 316
	Protein group(s): 2+ unique peptides, high confidence		
Range of peptide numbers	2 - 266	2 - 277	2 - 293
Range of peptide spectrum matches	2 - 1308	2 - 1748	2 - 1510
Molecular weight range (kDa)	5.02 - 3813.65	6.24 - 1010.45	6.24 - 3813.65
Coverage range	0.14 - 85.16	0.16 - 84.34	0.45 - 85.00
Proteins in replicate 1 to 3	3727		

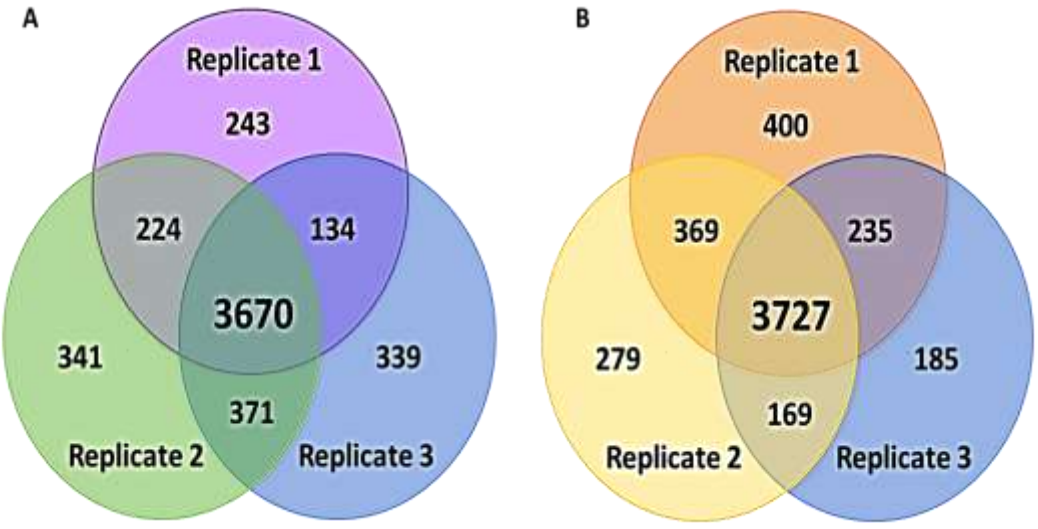

Figure 7.7: Venn diagram illustrating the overlap of proteins identified across replicates for A) CBT and B) CDK.

Histograms and summary statistics of coverage, molecular weight, peptides, PSMs and TMTs exhibited a high degree of similarity across all replicates for both experiments (Figure 7.8 and 7.9). Summary statistics and protein profiles of common housekeeping proteins (Figure 7.10 and 7.11) highlight the integrity of the data acquisition and processing.

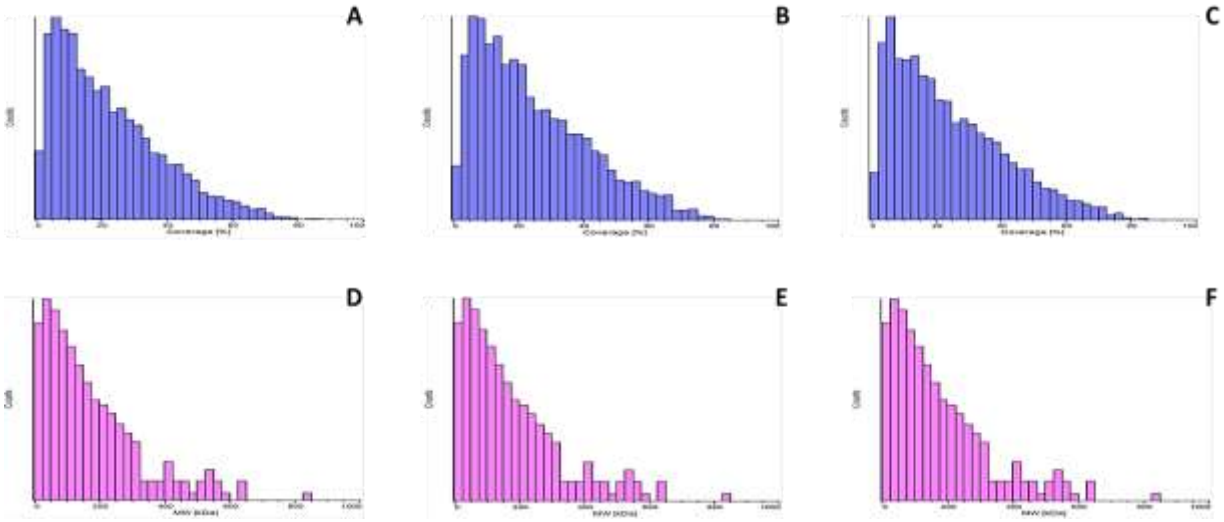

Figure 7.8: Histograms illustrating, A - C) Protein coverage (%) in CDK replicates 1 to 3 and D - F) Molecular weight (kDa) range of proteins in CDK replicates 1 to 3 (x-axis: counts).

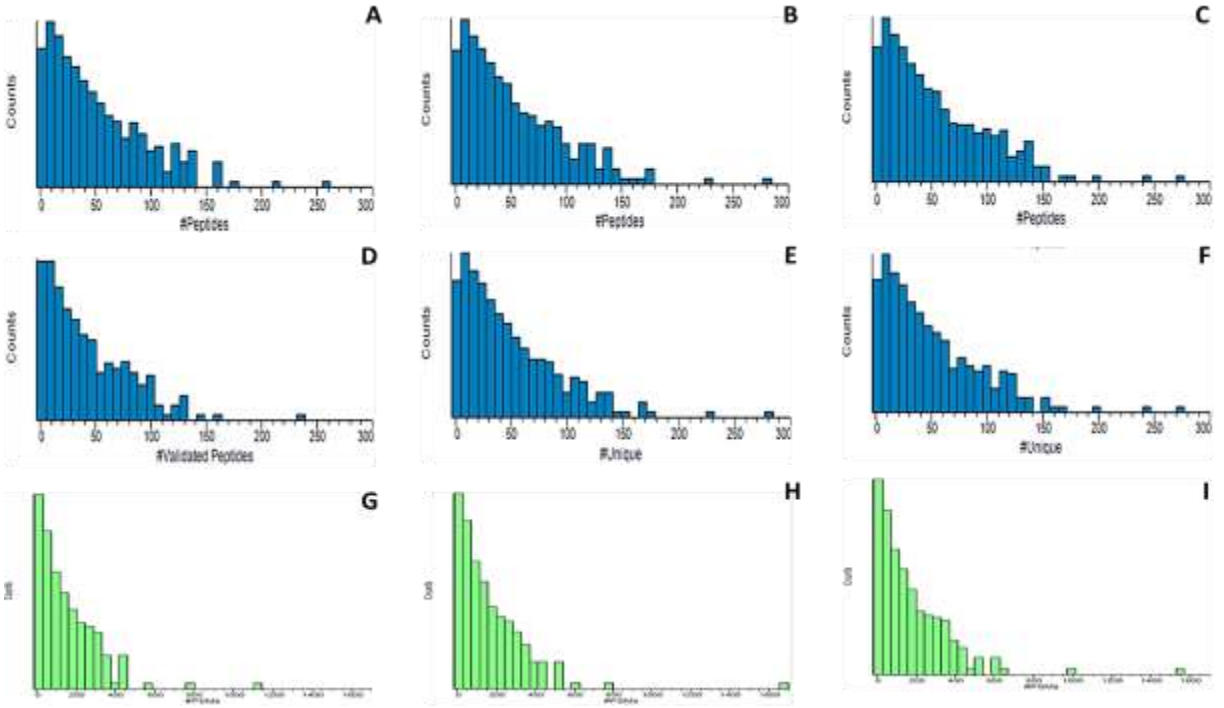

Figure 7.9: Histograms of CBT replicates 1 to 3 for A - C) Number of peptides, D - F) Number of unique peptides and G - I) Number of peptide-spectrum matches (x-axis: counts).

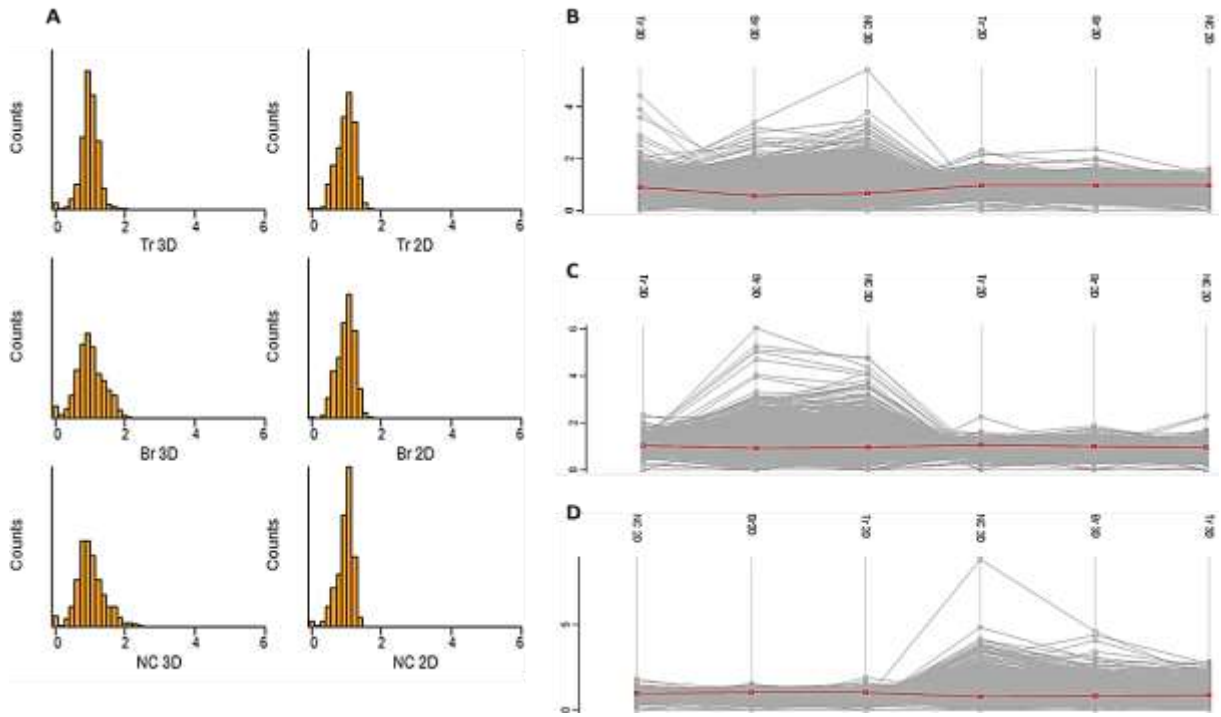

Figure 7.10: Tandem mass tag distribution and protein profiles. A) Histograms of CBT replicate 1 illustrating TMT distribution, B - D) Average protein profile for replicates 1 to 3. Red line on protein profiles: housekeeping protein glyceraldehyde-3-phosphate dehydrogenase (GAPDH; P04406).

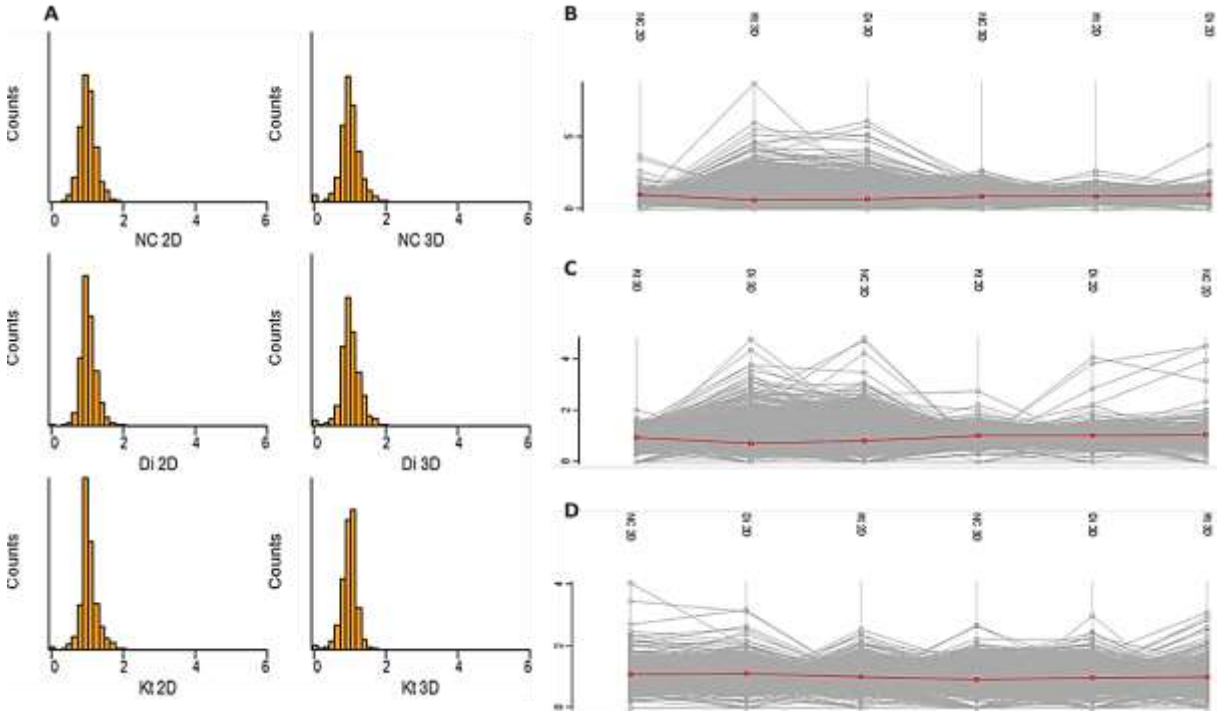

Figure 7.11: Tandem mass tag distribution and protein profiles. A) Histograms of CDK replicate 3 illustrating TMT distribution, B - D) Average protein profile for replicates 1 to 3. Red line on protein profiles: housekeeping protein glyceraldehyde-3-phosphate dehydrogenase (GAPDH; P04406).

7.2.1 Cell Type Comparisons: HepG2 monolayers versus HepG2 spheroids

Data supporting the notion that 3D cultures are the missing link for *in vitro* models to mimic *in vivo* counterparts is fragmentary. This is due to limited homogeneity in experimental designs. For the comparison between HepG2 monolayers and spheroids, proteins present in all six 6-plex experiments (n = 6 for monolayers and spheroids) was combined and an overlap of 3 421 proteins used to compare the two culture techniques. These comparisons are shown as volcano plots and hierarchical clustering (Figure 7.12).

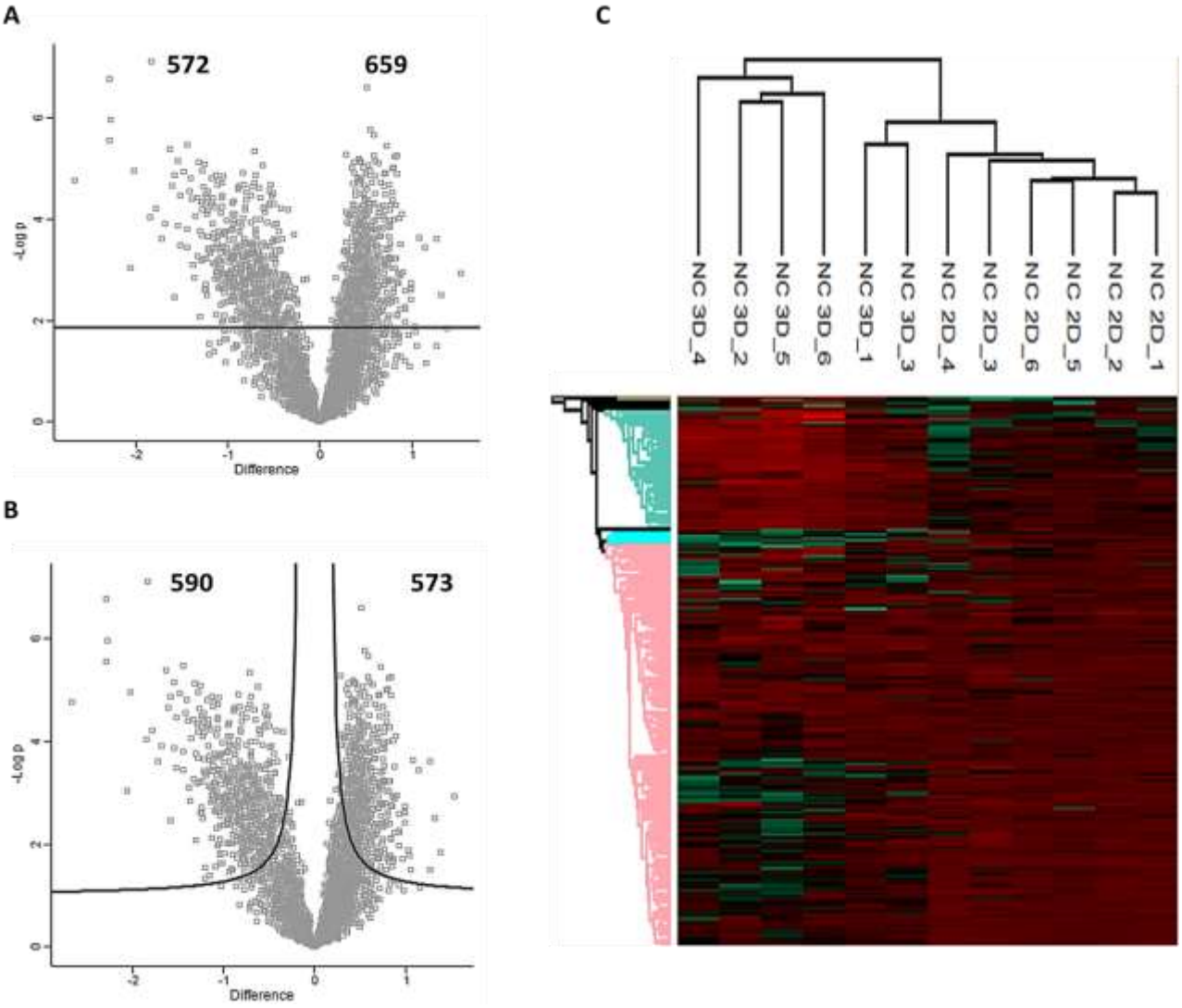

Figure 7.12: A) Volcano plot comparing HepG2 cell monolayers and spheroids using an FDR of 0.01 with contribution from the p-value only, B) Volcano plot comparing HepG2 cell monolayers and spheroids using an FDR of 0.01 with contributions from the p-value and mean difference and C) Hierarchical clustering of HepG2 cell monolayers and spheroids replicates which generally co-segregate base on the respective cell culture technique.

Others have observed that HepG2 cell spheroids have reduced proliferation, are capable of self-organization and differentiation. Altered gene expression of albumin, urea, xenobiotic transcription factors, phase I and II drug metabolizing enzymes and transporters becomes apparent between 14 and 28 days. Despite this, pronounced differences still exist, especially of phase I enzymes which remain poorly expressed compared to PHHs (Ramaiahgari SC, den Braver MW *et al.*, 2014). Time-dependent acquisition of liver-specific markers and enzymes at the transcriptome level would influence the dynamic proteome despite correlation between RNA and proteins remaining weak.

Enzymes and transporters involved in drug metabolism, so-called ADME proteins, are mostly integral membrane proteins or of low abundance which complicates detection and quantification (Vildhede A, Wisniewski JR *et al.*, 2015). ADME-associated proteins are described in numerous databases and include over 320 proteins from more than 50 protein families (Sun LZ, Ji ZL *et al.*, 2002). Approximately 700 genes code for proteins with relevance in the three groups of ADME proteins.

Vildhede *et al.* detected over 300 ADME proteins involved in phase I (CYP enzymes), phase II (UGT enzymes) and phase III (transport proteins) metabolism. Within these groups, CYP enzymes were the most abundant while drug transporters in the plasma membrane were generally less abundant, by one to two orders of magnitude, in isolated hepatocytes (Vildhede A, Wisniewski JR *et al.*, 2015). Detection of these proteins would be essential in order to establish the validity of any hepatotoxicity model.

An important question to consider within the context of this study was whether the proteins which were being augmented were more closely related with alterations in morphology or due to enhancement of metabolic capabilities. As was demonstrated in Chapter 5, there was an ability to increase expression of some of the liver-specific markers extracted from the data set. However, overall detection of less proteins in this experiments could have resulted in identification of fewer liver-specific markers as a result of their characteristic low abundance.

ECM proteins including fibronectin, vitronectin, laminin and tenascin, collagens (mainly type I), elastin and proteoglycans, can be secreted by most of the liver cell subpopulations. These proteins provide structural support to tissue and following binding to surface integrins, promote survival, continued expression of differentiated function and preservation of cell shape (Selden C, Khalil M *et al.*, 1999).

Proteins involved in cytoskeletal arrangements and ECM interactions varied in expression. Tropomyosins (P09493, P06753, P67936), filamins (P21333, O75369, Q14315), profilins (P35080, P07737), cofilins (P23528, Q9Y281), destrin (P60981), some kinesins (Q9NSK0, P33176, P52732, Q9NQT8) and catenins (P35221, P35222, O60716) were lower or similarly expressed in spheroids compared to monolayers.

Type I and II keratins (P13645, P05783, P08727, Q9C075, P35527, P04264, P35908, P13647, P04259, P05787), tubulins (P23258, P68363, Q9BQE3, P68366), integrins (P56199, P17301, P23229, P06756, P05556, P18084), fibronectin (P02751), vitronectin (P04004) and laminins (O15230, P07942, P11047) were increased in spheroids compared to monolayers.

The ECM and cytoskeletal arrangement dictate cell membrane polarity and organelle organisation of hepatocytes. The advantages naturally conferred by the ECM can be artificially introduced by collagen gels, laminin extracts and more complex combinations of matrices that re-establish cell membrane polarity and organelle organisation of hepatocytes (Selden C, Khalil M *et al.*, 1999). Addition of various matrices to the 3D culture microenvironment could further promote those ECM proteins that were not seen to be altered in this study.

Cell growth can be characterized by two extremes. Firstly, cells which grow exponentially have a short doubling time in monolayers. And secondly, cells can reach a dynamic equilibrium with a long doubling time in tissue-like conglomerates or spheroids. These physiological attributes appear to be gained spontaneously. HepG2/C3A cells cultured as monolayers for 5 days and as microgravity spheroids have been used to determine the changes that occur in the proteome as cells transition between exponential and equilibrium phases (Wrzesinski K, Rogowska-Wrzesinska A *et al.*, 2014).

Altered protein groups in spheroids include those governing cellular architecture and cytoskeletal rearrangement, metabolism (glucose, lipids, fatty acids), synthesis (urea, cholesterol), oxygen levels, cell growth, transport, ubiquitination and protein degradation. Cells grown as spheroids were capable of establishing an infrastructure which is perceived to enable better execution of advanced physiological function (Wrzesinski K, Rogowska-Wrzesinska A *et al.*, 2014). The data generated by Wrzensinski *et al.* was used to establish various relationships contributing to changes in the proteome following adaptation from 2D to 3D culture. This included linking of proteomic changes with observed morphological and physiological consequences. These observed relationships were assessed and correlations between these and the present study illustrated in Figure 7.13.

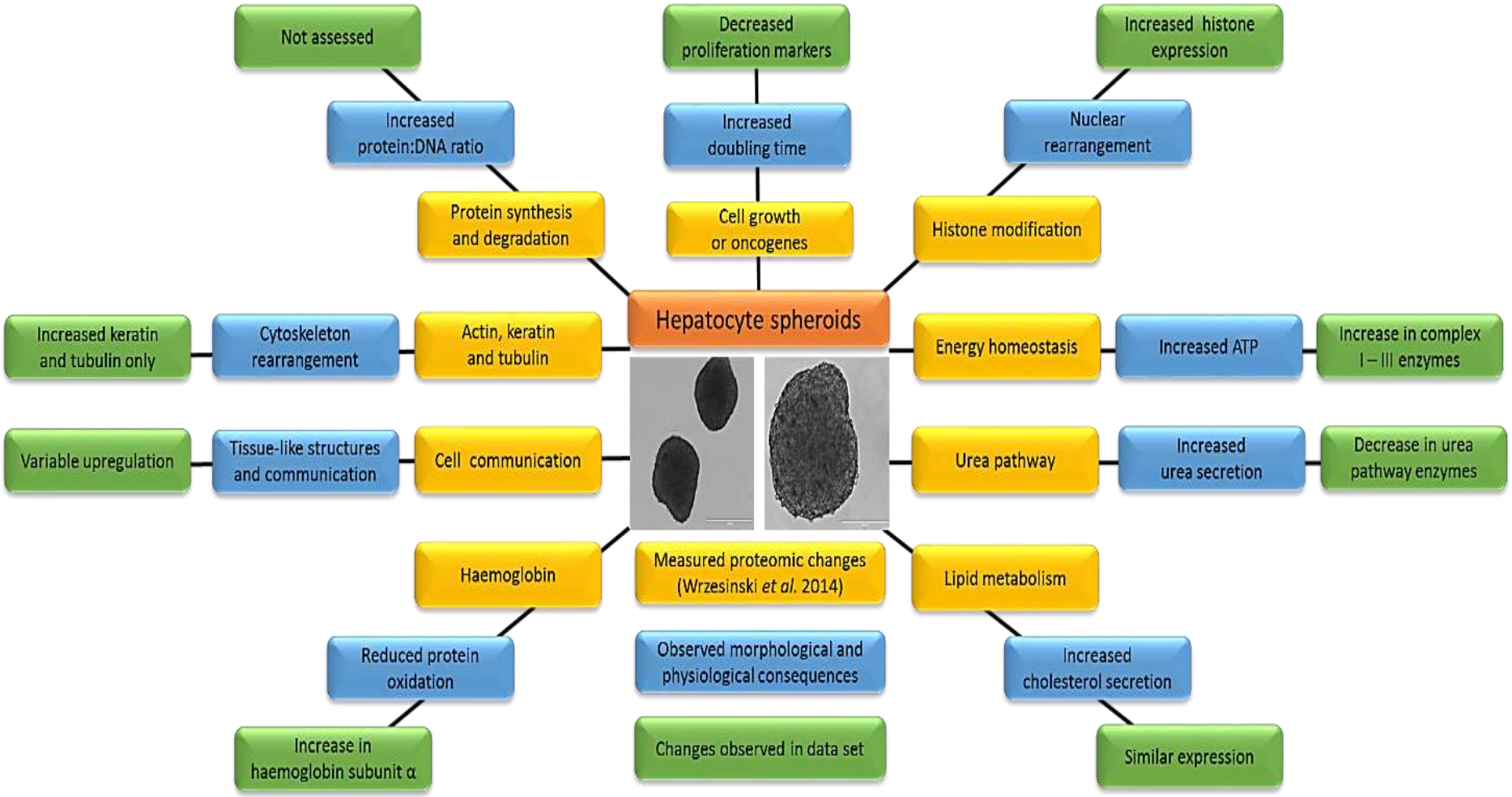

Figure 7.13: Relationship of the changes in the proteome following its adaptation from 2D to 3D culture with structural and physiological properties. Measured proteomic changes (yellow) with observed morphological and physiological consequences (blue) demonstrated by Wrzesinski et al. with the correlation or observations in this data set (green). (Reprinted from open access journal with permission, (Wrzesinski K, Rogowska-Wrzesinska A *et al.*, 2014))

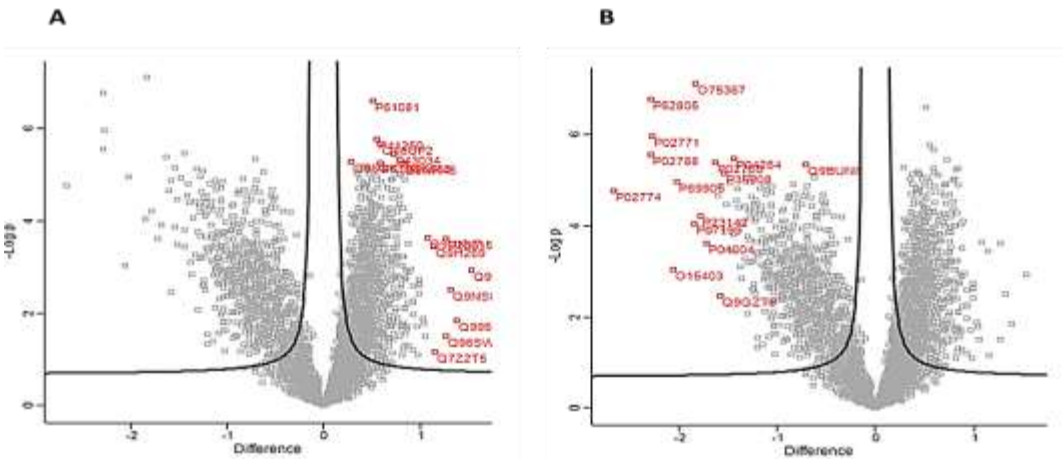

Figure 7.14: Proteins located at the extremes of volcano plots of HepG2 monolayers versus HepG2 spheroids which were investigated and described.

Table 7.10: Proteins located at extremes of volcano plots of HepG2 monolayers versus HepG2 spheroids (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with higher expression in HepG2 monolayers compared to HepG2 spheroids				
Vacuolar protein sorting-associated protein 35	Q96QK1	Retromer cargo-selective complex Lysosomal degradation pathway sorting	5.285	0.281
NEDD8-conjugating enzyme Ubc12	P61081	Pathway protein neddylation Involved in cell proliferation	6.604	0.516
Glycine--tRNA ligase	P41250	Catalyse attachment of glycine to tRNA cell regulation pathways	5.761	0.555
14-3-3 protein gamma	P61981	Adapter protein General / specialized signalling paths	5.245	0.585
Kinectin	Q86UP2	Kinesin-driven vesicle motility	5.662	0.589
Platelet-activating factor acetylhydrolase IB subunit alpha	P43034	Rho GTPases and actin polymerization Several dynein- and microtubule-dependent processes (G2/M transition)	5.443	0.724
Sialic acid synthase	Q9NR45	Lipopolysaccharide biosynthetic processes	5.229	0.818
High mobility group protein B2	P26583	Multifunctional protein with various roles in different cellular compartments	5.265	0.841
Protein CDV3 homolog	Q9UKY7	Implicated in cell proliferation	3.628	1.078
Vacuolar protein sorting-associated protein 16 homolog	Q9H269	Vesicle-mediated protein trafficking to lysosomal compartments Autophagy and protein transport	3.449	1.142
Protein cereblon	Q96SW2	Ubiquitination and proteasomal degradation of target proteins	1.495	1.264
Protein S100-P	P25815	Potential autocrine stimulation of cell proliferation	3.622	1.267
LanC-like protein 2	Q9NS86	Abscisic acid signalling pathway	2.508	1.317
Legumain	Q99538	Hydrolysis of asparaginyl bonds Plays a role in the regulation of cell proliferation via in EGFR degradation	1.840	1.375

Protein description	Main accession	Molecular function or biological process involvement	- Log (p-value)	Difference
Proteins with lower expression in HepG2 monolayers compared to HepG2 spheroids				
Vitamin D-binding protein	P02774	Vitamin D transport and storage	4.758	-2.666
Lactotransferrin	P02788	Iron binding transport proteins	5.549	-2.290
Histone H4	P62805	Core component of nucleosome Compact DNA into chromatin	6.764	-2.285
Alpha-fetoprotein	P02771	Bind copper, fatty acids, serum albumin Liver development and organ regeneration	5.966	-2.273
Monocarboxylate transporter 7	O15403	Proton-linked monocarboxylate (lactate, pyruvate) transporter	3.027	-2.062
Hemoglobin subunit alpha	P69905	Involved in oxygen transport	4.949	-2.025
Major centromere autoantigen B	P07199	Interacts with centromeric heterochromatin	4.045	-1.840
Core histone macro-H2A.1	O75367	Represses transcription in a subset of nucleosomes	7.119	-1.832
Fibulin-1	P23142	Fibronectin-containing matrix fibers Role in cell adhesion and migration within the ECM	4.206	-1.786
Vitronectin	P04004	Cell adhesion and spreading factor Interact with glycosaminoglycans and proteoglycan	3.620	-1.718
Alpha-2-HS-glycoprotein	P02765	Promotes endocytosis, possesses opsonic properties	5.379	-1.629
Keratin. type II cytoskeletal 2 epidermal	P35908	structural constituent of cytoskeleton	5.153	-1.546
Keratin. type II cytoskeletal 1	P04264	May regulate the activity of kinases	5.475	-1.444
Derlin-1	Q9BUN8	Functional component of endoplasmic reticulum-associated degradation for misfolded luminal proteins	5.343	-0.704

Trypsin proteolytically degrades essential ECM proteins and activates several protease-activated receptors (PAR 1-4) which can effect gene expression. The regularity with which trypsinization (for passaging confluent cells) is conducted can hinder cells from regaining the natural morphology, physiological function and signal transduction. Wrzesinski *et al.* demonstrated that several physiological functions of HepG2/C3 cells are suppressed following trypsinization with the reestablishment of these features initially occurring at similar rates (approximately 5 days) in monolayer and spheroid cultures. In spheroid cultures, re-acquisition of features can occur for up to 18 days. It is postulated that in monolayers, cellular resources are focussed on repair and proliferation and forego differentiation processes. While in spheroids, cells can re-establish and then re-differentiate and specialize termed contact functional activation, to represent the *in vivo* living organism (Wrzesinski K and Fey SJ, 2013).

Regularity of trypsinization and the time prior to trypsin exposure greatly influences the hepatic phenotype. In this study, HepG2 cells for both monolayer and spheroid cultures were trypsinized at the same time from the same flask and seeded (Day 0) followed by culturing for 4 days prior to exposure. This would have allowed the re-establishment of physiological function reduced by trypsin. However, it may not have been sufficient for spheroids to further re-differentiate and promote functional activation in the presence of drug-induced stress.

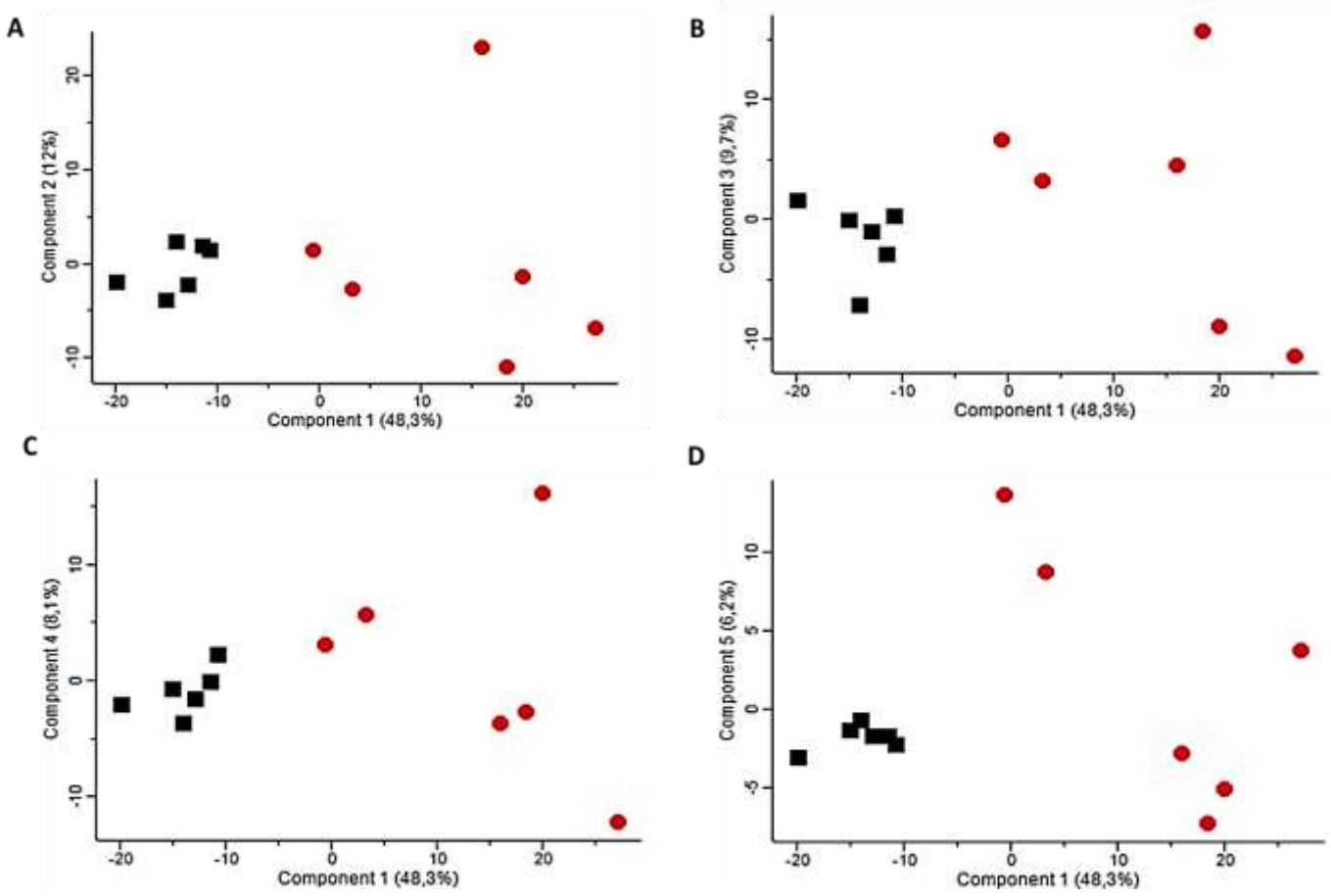

Figure 7.15: Principal component analysis of HepG2 cell monolayers and spheroids. A) Component 1 versus Component 2, B) Component 1 versus Component 3, C) Component 1 versus Component 3 and D) Component 1 versus Component 4. HepG2 cells monolayers: black square (n = 6) and HepG2 spheroids: red circle (n = 6).

PCA clustered all monolayers cultures distinctly in Component 1 (48.53%) and Component 2 (12.0%) of principal component space (Figure 7.15A). A total of 11 components were distinguished which contributed to between 4.6 - 1.1% of the variance in Components 6 to 11. HepG2 spheroids were spatially resolved from monolayers, however, there was no tight clustering of samples. This could reveal the potential heterogeneity with which the HepG2 cell line undergoes development and maturation when cultured as spheroids.

7.2.2 Proteomic Profiling of Cytotoxicity: Bromfenac

Bromfenac, a benzene acetic acid derivative containing a carboxylic acid moiety, is a potent peripherally acting non-narcotic analgesic nonsteroidal anti-inflammatory drug. Following market approval in 1997, Bromfenac was available for 114 months prior to drug withdrawal. Prolonged Bromfenac use, for more than the ten days as recommended by manufacturers, was implicated in severe hepatotoxicity. Elevated transaminases, of up to three times the ULN, were observed in approximately 15% of patients (Hunter EB, Johnston PE *et al.*, 1999; Moses PL, Schroeder B *et al.*, 1999). Human urine analysis has detected five reactive metabolites including a cyclic amide and two pairs of polar diastereoisomeric glucuronide metabolites (Walgren JL, Mitchell MD *et al.*, 2005).

Histological features of severe Bromfenac hepatotoxicity included multilobular necrosis or collapse, lobular inflammation, periportal inflammation, steatosis and nodular regeneration (Fontana RJ, McCashland TM *et al.*, 1999). Reactive metabolites of Bromfenac potentially overwhelm detoxification pathways and bind cellular proteins through sulphydryl modification to elicit cellular damage (Williams DP and Park BK, 2003). Volcano plots of identified and quantitated proteins from Bromfenac exposed versus controls (Figure 7.16), did not reveal any proteins which differed significantly in either monolayers or spheroids.

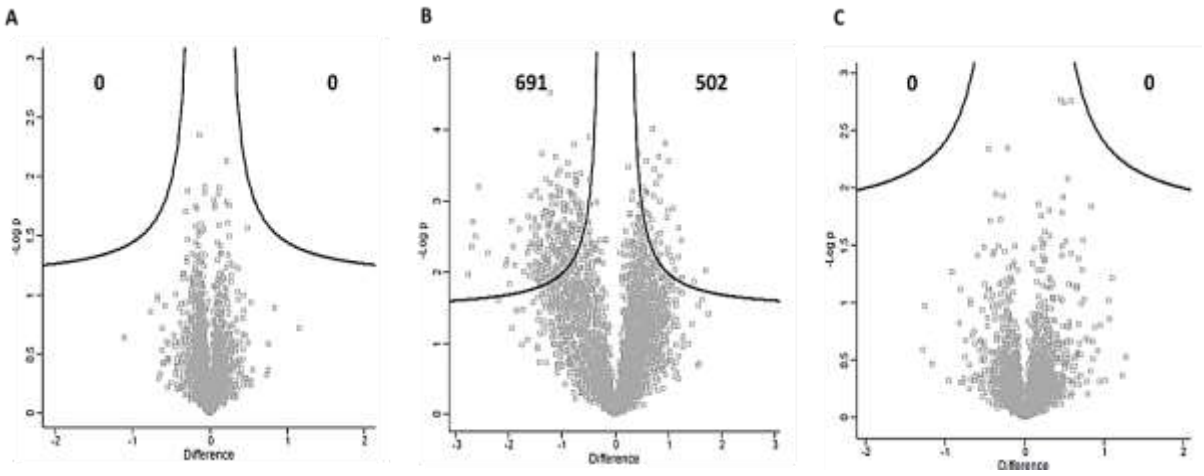

Figure 7.16: Volcano plot comparing HepG2 cells using an FDR of 0.01 with contributions from the p-value and mean difference. A) Negative control monolayers versus Bromfenac monolayers, B) Negative control monolayers versus Bromfenac spheroids and C) Negative control spheroids versus Bromfenac spheroids.

The differences observed in plots of negative control monolayers versus Bromfenac-exposed spheroids (Figure 7.16B) was primarily accounted for by the variance introduced by the culturing method. Slightly more proteins were reduced in Bromfenac-exposed spheroids than for the monolayer versus spheroid cultures. However, as monolayers and spheroids exposed to Bromfenac did not present statistically significant differences compared to untreated controls, protein profiles were used to identify more subtle variances (Figure 7.17 and Table 7.11).

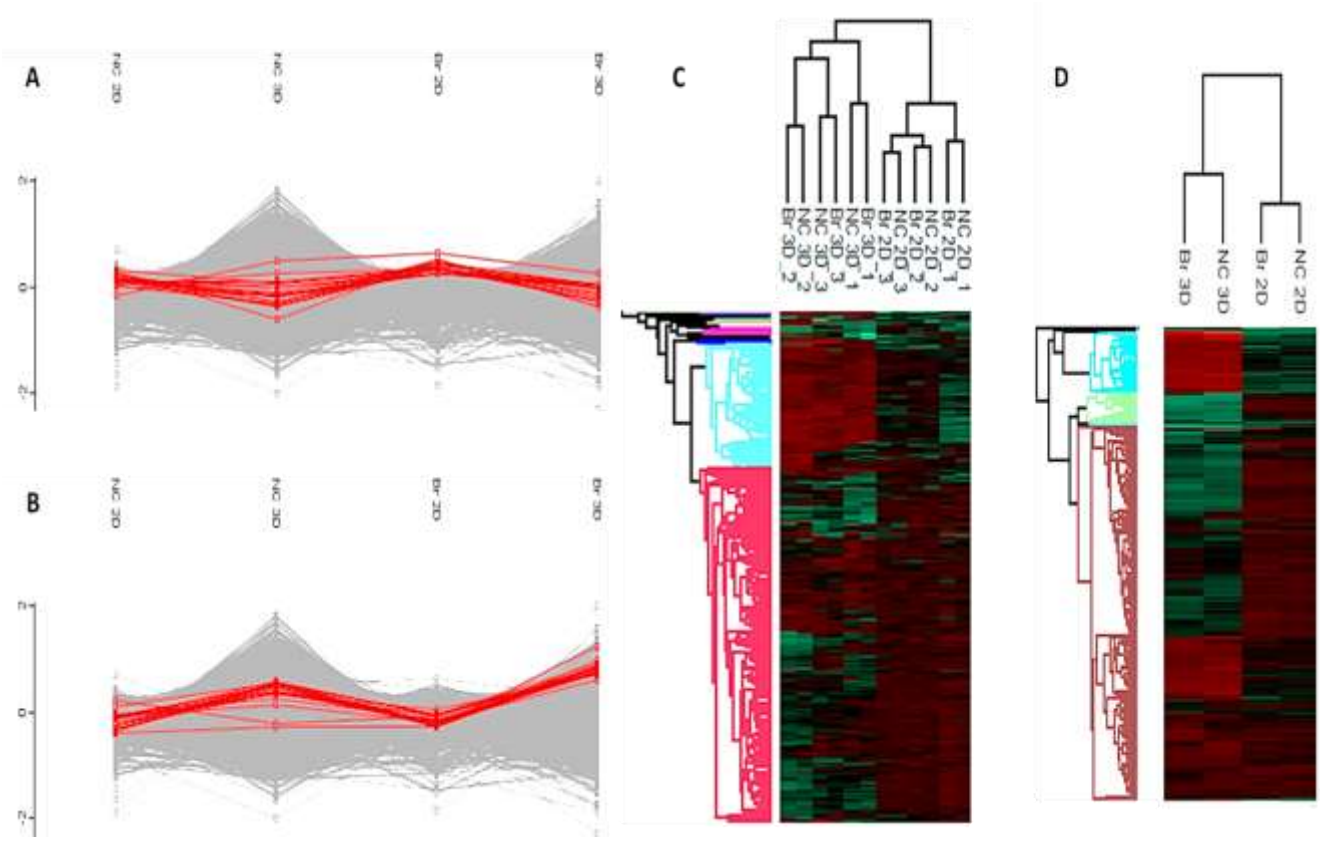

Figure 7.17: Protein cluster profiles used to identify proteins peaking in A) HepG2 monolayers exposed to Bromfenac and B) HepG2 spheroids exposed to Bromfenac C) Hierarchical clustering of ten column clusters for replicates illustrating co-segregation of monolayers versus spheroids with their respective sample exposed to Bromfenac and D) Hierarchical clustering using the average of replicates.

Table 7.11: Top proteins which peaked in HepG2 monolayers or HepG2 spheroids exposed to Bromfenac showing minimal change in respective negative controls (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in HepG2 monolayers exposed to Bromfenac		
Sulfotransferase 1C2	O00338	Utilizes 3'-phospho-5'-adenylyl sulfate as sulfonate donor to catalyse the sulfate conjugation of drugs, xenobiotic compounds
Eukaryotic elongation factor 2 kinase	O00418	Threonine kinase that regulates protein synthesis by controlling the rate of peptide chain elongation
LanC-like protein 1	O43813	May play a role in EGFR kinase substrate 8 signalling Binds glutathione
Superoxide dismutase [Mn], mitochondrial	P04179	Destroys superoxide anion radicals normally produced within the cells and which are toxic to biological systems
Tissue factor pathway inhibitor	P10646	Antithrombotic action and also the ability to associate with lipoproteins in plasma
60S ribosomal protein L35a	P18077	Plays a role in 60S ribosomal subunit formation
RNA-binding motif, single-stranded-interactor 1	P29558	Single-stranded DNA binding protein that interacts with the region upstream of the MYC gene
DNA ligase 3	P49916	Correct defective DNA strand-break repair and sister chromatid exchange following ionizing radiation and alkylating agents
Hydroxymethylglutaryl-CoA synthase, cytoplasmic	Q01581	Condenses acetyl-CoA with acetoacetyl-CoA to form HMG-CoA (HMG-CoA reductase pathway)
Syntaxin-5	Q13190	Mediates endoplasmic reticulum to Golgi transport
Proteins peaking in HepG2 spheroids exposed to Bromfenac		
Serine palmitoyl transferase 1	O15269	Protein is involved in the pathway sphingolipid (fatty acid derivatives of sphingosine) metabolism
Low-density lipoprotein receptor	P01130	Binds LDL, the major cholesterol-carrying lipoprotein of plasma, and transports it into cells by endocytosis
Propionyl-CoA carboxylase alpha chain, mitochondrial	P05165	Sub-pathway that synthesizes succinyl-CoA from propanoyl-CoA (metabolic intermediary metabolism)
Solute carrier family 2, facilitated glucose transporter member 2	P11168	Mediates the bidirectional transfer of glucose across the plasma membrane of hepatocytes
Myosin-9	P35579	Role in cytokinesis, cell shape, and specialized functions such as secretion and capping
Myosin-10	P35580	Role in cytokinesis, cell shape, and specialized functions such as secretion and capping
Signal peptidase complex catalytic subunit SEC11A	P67812	Microsomal signal peptidase complex which removes signal peptides from nascent proteins as they are translocated
Malectin	Q14165	Carbohydrate-binding protein Role in early steps of protein N-glycosylation
E3 ubiquitin-protein ligase synoviolin	Q86TM6	Involved in ubiquitin-dependent degradation of misfolded endoplasmic reticulum proteins
Complex I assembly factor TMEM126B, mitochondrial	Q8IUX1	Chaperone protein involved in the assembly of the mitochondrial NADH:ubiquinone oxidoreductase complex

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in HepG2 monolayers and spheroids exposed to Bromfenac		
RNA 3'-terminal phosphate cyclase	O00442	Catalyses the conversion of 3'-phosphate to a 2',3'-cyclic phosphodiester at the end of RNA
Solute carrier family 22 member 3	O75751	Mediates potential-dependent transport of organic cations
Regulator of chromosome condensation	P18754	Involved in the regulation of onset of chromosome condensation in the S phase
Cadherin-2	P19022	Calcium-dependent cell adhesion proteins
Sodium- and chloride-dependent glycine transporter 1	P48067	Terminates the action of glycine by its high affinity sodium-dependent reuptake
Lanosterol synthase	P48449	Catalyses the cyclization of (S)-2,3 oxidosqualene to lanosterol, a reaction that forms the sterol nucleus
Sodium-dependent phosphate transporter 2	Q08357	Plays a fundamental housekeeping role in phosphate transport
Probable ATP-dependent RNA helicase DDX10	Q13206	Putative ATP-dependent RNA helicase
Sequestosome-1	Q13501	Autophagy receptor
39S ribosomal protein L51	Q4U2R6	Structural component of ribosomes: mitochondrial translation

Following exposure to Bromfenac in monolayer cultures, proteins which peaked included sulfotransferase 1C2 (O00338), superoxide dismutase (P04179) and lanC-like protein 1 (O43813). These proteins all play a role in phase II detoxification and can be linked to detoxification pathways involving glutathione. Proteins implicated in detoxification were not as readily observed in spheroids exposed to Bromfenac. This could be due to the fact that the changes induced by the different culturing technique were greater than the changes induced by drug exposure, thereby masking potential toxicity profiles.

In Chapter 6, an *in vitro* mechanistic toxicity screening model, using cultured HepG2 monolayers was conducted to elucidate hepatotoxic mechanisms. Of the 6 assays assess, some proteins correlating to each process were extracted. Little variance was observed in extracted proteins between monolayer and spheroid cultures with no differential expression following Bromfenac exposure. Clinically relevant elevated transaminases (ALT/AST) have been observed in 15% of patients following Bromfenac use. ALT is more liver specific than AST but can be increased without clinical consequence (Yang X, Salminen WF *et al.*, 2012). Cytoplasmic and mitochondrial AST were decreased and increased respectively in spheroids cultures respectively and showed no alteration following Bromfenac exposure. As transaminases are usually released into the medium, it remains possible that levels were increased but not observed due to discarding the medium during samples collection.

Additionally, five reactive drug metabolites are associated with Bromfenac toxicity (Walgren JL, Mitchell MD *et al.*, 2005). Compared to PHHs, HepG2 cells in both monolayers and spheroids (Chapter 5), express inadequate levels of metabolic enzymes. This suggests that both culture formats are unable to sufficiently transform Bromfenac to reactive metabolites. The inability to detect hepatotoxicity with these models could also be related to limiting drug exposure. The typical daily dose of Bromfenac associated with idiosyncratic hepatotoxicity (25 - 50 mg/day) was used to calculated the C_{max} of Bromfenac from a dose of 50 mg (25 µM). However, exposure in the experimental design was limited to one repeated exposure over a period of 6 days. This could mean that the possibility exists that the time frame of exposure was insufficient to induce hepatotoxicity.

7.2.3 Proteomic Profiling of Cytotoxicity: Ketoconazole

Oral formulations of Ketoconazole were originally given black box warnings (reasonable evidence of an association with a serious hazard) primarily due to its potent inhibitory effect on CYP3A4 with the FDA and EMA now limiting oral use due to severe hepatic toxicity. Extensive metabolism is dependent on hepatic microsomal enzymes facilitating biotransformation reactions. It undergoes N-deacetylation to form N-deacetyl-ketoconazole (DAK) after which a series of flavin-containing monooxygenase (FMO)-dependent reactions result in N-hydroxy DAK which finally forms a reactive dialdehyde (Walgren JL, Mitchell MD *et al.*, 2005). In rat hepatocytes, DAK has demonstrated a time and dose-related cytotoxicity which is more potent than that of Ketoconazole (Rodriguez RJ and Acosta D, 1997).

Human FMO enzymes, FMO1 and FMO3 have been implicated in catalysing N-hydroxylation of DAK with the enzymes converting Ketoconazole to DAK long remaining unknown. The two esterases long thought to be implicated in mediating this reaction were carboxylesterase (CES) and arylacetamide deacetylase (AADAC). A recent study by Fukami *et al.* elucidated that AADAC (P22760) is the primary enzyme responsible for Ketoconazole hydrolysis into DAK with FOM1 (expressed only in the kidney) and FOM3 required for further metabolism (Fukami T, Iida A *et al.*, 2016).

Only four proteins, identified from volcano plots (Figure 7.18) were reduced (0.5 fold) in HepG2 monolayers exposed to Ketoconazole. These included insulin receptor substrate 1 (P35568), paternally-expressed gene 3 protein (Q9GZU2) involved in cooperative induction of apoptosis with E3 ubiquitin-protein ligase which mediates ubiquitination, proliferating cell nuclear antigen (P12004) controlling DNA replication and heterogeneous nuclear ribonucleoprotein M (P52272) associated with RNA splicing. Similar reductions were seen in spheroid cultures, but subtle changes appear masked by the greater changes induced by culturing as spheroids.

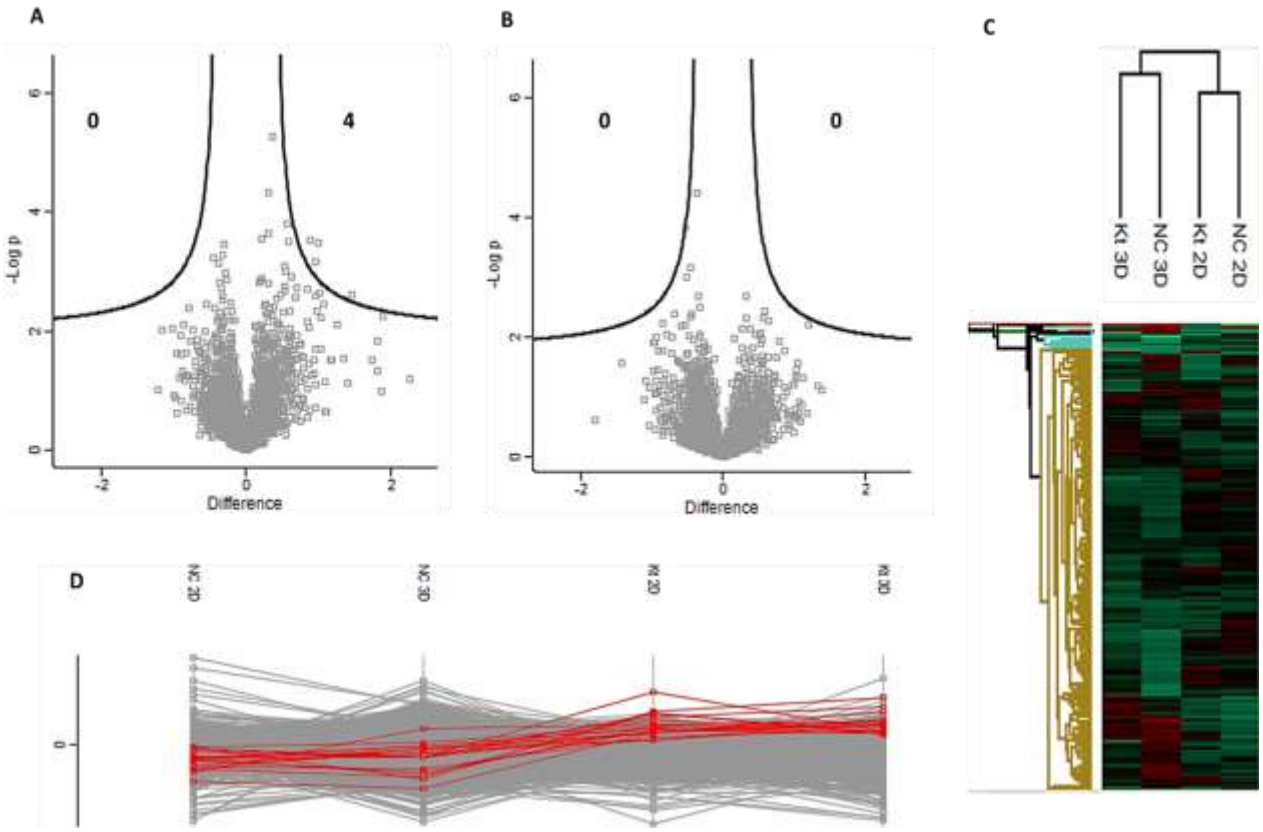

Figure 7.18: Volcano plot comparing HepG2 cells using an FDR of 0.01 with contributions from the p-value and mean difference. A) Negative control monolayers versus Ketoconazole monolayers, B) Negative control spheroids versus Ketoconazole spheroids, C) Hierarchical clustering and D) Protein profiles used to identify proteins increasing in monolayers and spheroids exposed to Ketoconazole.

However, none of these proteins provided insight into the mechanism of cell death induced by Ketoconazole. The ability to detect proteins which peaked in one experimental condition but remained constant throughout the others conditions was restricted by the limited range of abundance. Therefore, when grouping 20 proteins, based on Euclidean distance, there was no tight grouping in the proteins extracted (Table 7.12).

Table 7.12: Top proteins which peaked in monolayers or spheroids exposed to Ketoconazole showing minimal change in respective negative controls (TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in HepG2 monolayers exposed to Ketoconazole		
Interferon-related developmental regulator	O00458	Role in regulating gene activity in the proliferative and/or differentiation pathways induced by NGF
Annexin A3	P12429	Inhibitor of phospholipase A2 (release fatty acids from glycerol)
Histone H1x	Q92522	Condensation of nucleosome chains into higher-order structures
Transgelin	Q01995	Actin cross-linking/gelling protein Involved in calcium interactions and contractile properties of the cell that may contribute to replicative senescence
60S ribosomal protein L23a	P62750	Protein binds to a specific region on the 26S rRNA Involved in biological process associated with cell proliferation
Proteins peaking in HepG2 spheroids exposed to Ketoconazole		
Protein CYR61	O00622	Promotes cell proliferation, chemotaxis, angiogenesis and cell adhesion
CDP-diacylglycerol-inositol 3-phosphatidyltransferase	O14735	Catalyses the biosynthesis of phosphatidylinositol
H/ACA ribonucleoprotein complex subunit 4	O60832	Required for ribosome biogenesis and telomere maintenance Promotes cell to cell and cell to substratum adhesion, increases the cell proliferation rate
Tyrosyl-DNA phosphodiesterase 2	O95551	DNA repair enzyme that can remove a variety of covalent adducts from DNA Acts as a regulator of ribosome biogenesis following stress
Heterogeneous nuclear ribonucleoproteins C1/C2	P07910	Binds pre-mRNA and nucleates the assembly of 40S hnRNP particles
Proteins peaking in HepG2 monolayers and spheroids exposed to Ketoconazole		
Lanosterol 14-alpha demethylase	Q16850	Catalyses C14-demethylation of lanosterol
Differentiation factor 15	Q99988	Cytokine activity for cell signalling and development
Protein CYR61	O00622	Promotes cell proliferation, chemotaxis, angiogenesis and cell adhesion
Zinc transporter ZIP4	Q6P5W5	Plays an important role in cellular zinc homeostasis as a zinc transporter
Epithelial cell adhesion molecule	P16422	May act as a physical homophilic interaction molecule
Occludin	Q16625	May play a role in formation and regulation of tight junctions Induce adhesion when expressed in cells lacking tight junctions
Poliovirus receptor	P15151	Mediates NK cell adhesion and triggers NK cell effector functions
CKLF-like MARVEL transmembrane domain-containing protein 6	Q9NX76	Protein involved in chemotaxis
RRP15-like protein	Q9Y3B9	Maturation of rRNA
Ninjurin-1	Q92982	Homophilic cell adhesion molecule that promotes axonal growth

AADAC is required to hydrolyse Ketoconazole and to trigger downstream effects associated with hepatocellular toxicity (Fukami T, Iida A *et al.*, 2016). AADAC was not identified in any of the replicates, even prior to filtering for protein overlap. Successive oxidation attacks, require the flavin-containing monooxygenase system, to produce the more toxic dialdehyde. Only one FMO, dimethylaniline monooxygenase [N-oxide-forming] 5 (FMO5, P49326), was identified in the data set. FMO5, has only shown marginal activity compared to FMO3 and FMO1 which are essential in the liver and kidney respectively. Hepatotoxicity was induced by Ketoconazole at concentrations above 10 µM using the multiplexed hepatotoxicity model. This could be due to direct toxicity of Ketoconazole which is less potent than DAK and other reactive metabolites.

Inability to elucidate changing proteomic profiles of Ketoconazole-induced toxicity could be due to inadequate metabolism required to generate the known reactive metabolites. Furthermore, the typical daily dose of Ketoconazole associated with idiosyncratic hepatotoxicity is 200 mg which is equivalent to a plasma concentration range of 8 - 19 µM (Rodriguez RJ and Acosta D, 1997). Having only exposed to 20 µM over a period of 6 days could be an additional factor in the inability to elicit a hepatotoxic effect detectable at the proteome level.

7.2.4 Proteomic Profiling of Cytotoxicity: Troglitazone

Troglitazone, is an anti-diabetic, which was used to decrease resistance to insulin via peroxisome proliferator-activated receptor (PPAR) agonism. It is an inducer and selective substrate of the CYP3A enzyme family. Troglitazone is composed of a thiazolidinedione ring which is common to the class of glitazone drugs, and a chromane ring which is unique to Troglitazone. Oxidative bioactivation of both moieties to reactive metabolites requires elimination as thioether or thioester conjugates of glutathione (Kassahun K, Pearson PG *et al.*, 2001; Park BK, Kitteringham NR *et al.*, 2005).

The two ring structures provide two distinct bioactivation pathways for Troglitazone. The chromane ring can undergo oxidation via multiple non-specific P450 enzymes to reactive o-quinone methide. The thiazolidinedione ring undergoes an oxidative cleavage via CYP3A4 to α-ketoisocyante via sulfenic acid intermediates (Kassahun K, Pearson PG *et al.*, 2001; Walgren JL, Mitchell MD *et al.*, 2005). Metabolism of Troglitazone results in five possible GSH conjugates which are implicated in the severe hepatotoxicity (Kassahun K, Pearson PG *et al.*, 2001).

Biotransformed, chemically reactive metabolites may be responsible for Troglitazone-induced injury. However, direct toxicity resulting in mitochondrial dysfunction, alteration of mitochondrial permeability, apoptosis, PPAR-dependent steatosis and cholestasis due to inhibition of bile salt export pumps has also been reported (Masubuchi Y, 2006).

Hierarchical clustering was done using averages of the replicates (Figure 7.19). Generating ten row clusters produced clusters ranging from 1 to 2 820 proteins per cluster. In contrast to Bromfenac and Ketoconazole which co-segregated according to cell culture technique, when analysing the Troglitazone data, the main node of negative control spheroids encompassed all other samples. HepG2 cells cultured as monolayers remained co-segregated with spheroids exposed to Troglitazone adjoining the subsequent node. The exact reason behind this inconsistency is unknown but detailed investigation of cluster variance between samples could aid in elucidating this.

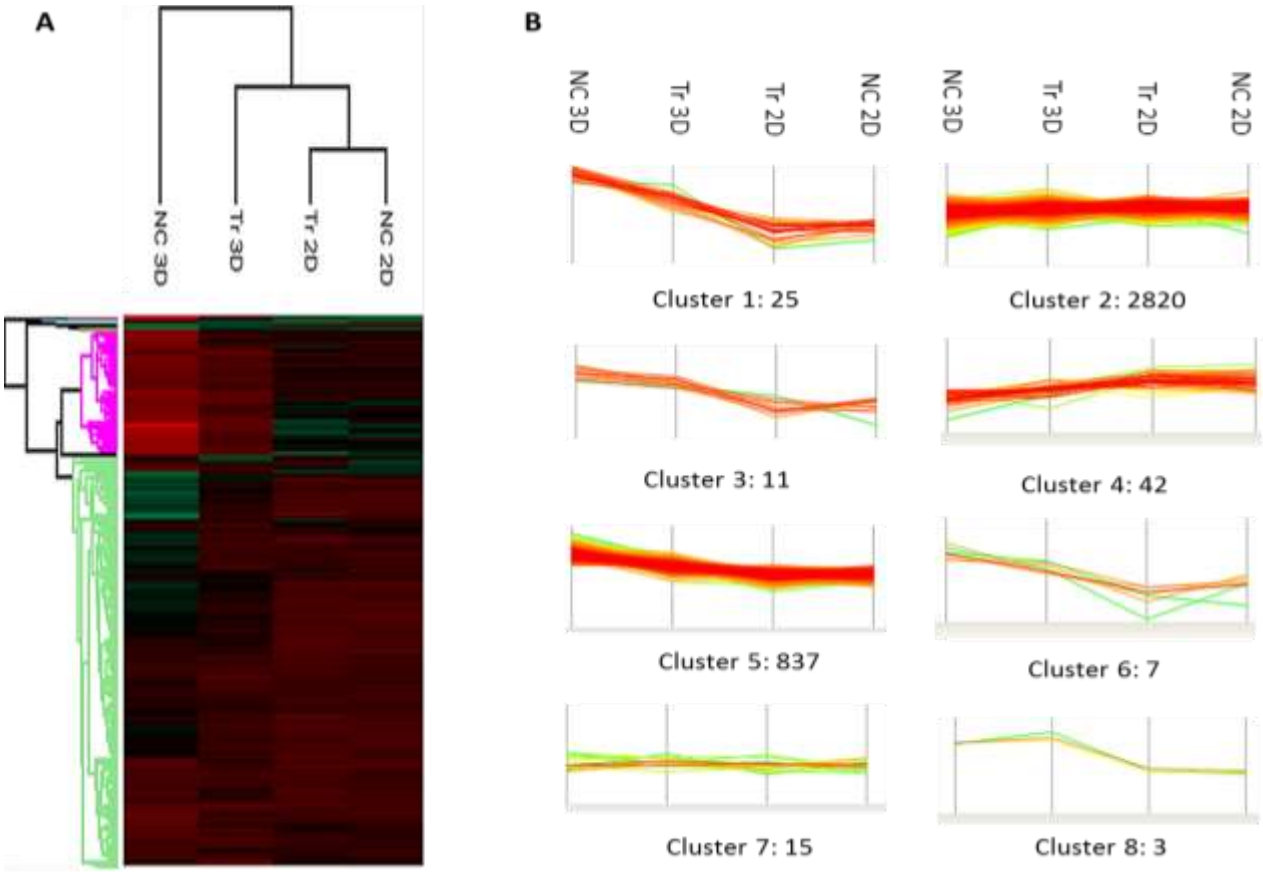

Figure 7.19: A) Hierarchical clustering of replicate averages and B) Protein profiles produced when defining 10 row clusters with the corresponding number of proteins. Clusters 9 and 10 contained three and one protein respectively.

Volcano plots distinguished 10 differentially regulated proteins in exposed spheroids (Figure 7.20) which were described (Table 7.13). Additionally, proteins peaking in monolayers and spheroids exposed to Troglitazone were identified.

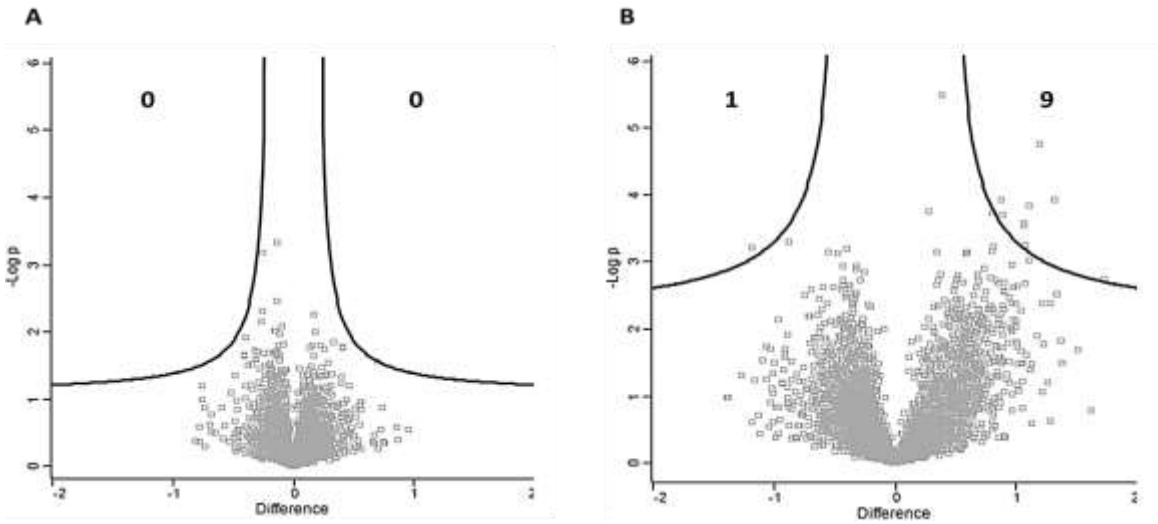

Figure 7.20: Volcano plot comparing HepG2 cells using an FDR of 0.01 with contributions from the p-value and mean difference. A) Negative control monolayers versus Troglitazone monolayers and B) Negative control spheroids versus Troglitazone spheroids.

Table 7.13: Proteins distinguished as significant in volcano plots of controls versus Troglitazone exposed spheroids and proteins which peaked in exposed monolayers and spheroids
(TheUniProtConsortium, 2014)

Protein description	Main accession	Molecular function or biological process involvement
Higher expression in HepG2 control spheroids compared to spheroids exposed to Troglitazone		
Vitamin D-binding protein	P02774	Involved in vitamin D transport and storage
Keratin, type II cytoskeletal 2 epidermal	P35908	Structural constituent of cytoskeleton
Keratin, type I cytoskeletal 9, 10	P35527 P13645	Structural constituent of epidermis
Keratin, type II cytoskeletal 1	P04264	Regulate the activity of kinases such as PKC and SRC
Keratin, type II cytoskeletal 6B	P04259	Structural constituent of cytoskeleton
Cytochrome b-c1 complex subunit 7	P14927	Ubiquinol-cytochrome c reductase complex (complex III), part of mitochondrial respiratory chain
CDP-diacylglycerol--inositol 3-phosphatidyltransferase	O14735	Catalyses the biosynthesis of phosphatidylinositol
Epoxide hydrolase 1	P07099	Biotransformation enzyme that catalyses the hydrolysis of arene and aliphatic epoxides to less reactive and more water soluble dihydrodiols
Peptidyl-prolyl cis-trans isomerase FKBP7	Q9Y680	Accelerate the folding of proteins during protein synthesis
Higher expression in HepG2 spheroids exposed to Troglitazone compared to control spheroids		
Vacuolar protein sorting-associated protein 16 homolog	Q9H269	Plays a role in vesicle-mediated protein trafficking to lysosomal compartments including endocytic membrane transport and autophagic pathways

Protein description	Main accession	Molecular function or biological process involvement
Proteins peaking in both HepG2 monolayers and spheroids exposed to Troglitazone		
Aldo-keto reductase family 1 member B10	O60218	Acts as all-trans-retinaldehyde reductase efficiently reduce aliphatic and aromatic aldehydes
rRNA methyltransferase 3, mitochondrial	Q9HC36	Catalyses the formation of 2'-O-methylguanosine in the 16S mitochondrial large subunit ribosomal RNA
Probable tRNA N6-adenosine threonyl carbamoyl transferase	Q9NPF4	Required for formation of threonylcarbamoyl groups on adenosine
Aldo-keto reductase family 1 member C1	Q04828	Converts progesterone to its inactive form Role in monitoring intrahepatic bile acid concentration
Aldo-keto reductase family 1 member C2	P52895	Convert steroid hormones high bile-binding ability
High mobility group nucleosome-binding domain- protein 3	Q15651	Binds to nucleosomes, regulating chromatin structure and consequently, chromatin-dependent processes Regulates the expression of the glucose transporter SLC2A2
F-box-like/WD repeat-containing protein TBL1XR1	Q9BZK7	Recruitment of the ubiquitin/19S proteasome complex to nuclear receptor-regulated transcription units
Very-long-chain 3-oxoacyl-CoA reductase	Q53GQ0	Catalyzes the second of the four reactions of the long-chain fatty acids elongation cycle
Striatin-interacting protein 1	Q5VSL9	Regulation of cell morphology and cytoskeletal organization
Catechol O-methyltransferase	P21964	Inactivation of catecholamine neurotransmitters and catechol hormone

Control spheroids, cultured for 10 days were capable of altering the HepG2 proteome. In contrast, spheroids which were exposed to Troglitazone were only grown for 4 days prior to initial drug exposure. This could highlight why much of the protein variance between spheroids in exposed versus non-exposed cultures were associated with structural functions (Table 7.13). Following exposure to a toxic compound, cells could switch to processes required for cell survival as opposed to control spheroids capable of continued maturation throughout the duration of culture.

Troglitazone, undergoes metabolism to reactive o-quinone methide and α-ketoisocyante via multiple nonspecific P450 enzymes and CYP3A4 respectively (Kassahun K, Pearson PG *et al.*, 2001; Walgren JL, Mitchell MD *et al.*, 2005). This proteomic dataset contained few CYP enzymes (CYP20A1, CYP27A1, CYP2S1, CYP2W1, CYP4F11 and CYP51A1), which are not considered of the major metabolising enzymes. Importantly, CYP3A4 was absent from this dataset which provides an explanation for the inability of this HepG2 model system to bring about Troglitazone metabolite-associated toxicity. Despite exposing to 15 µM twice over a period of 6 days, no proteomic profile indicative of Troglitazone induced hepatotoxicity was detected which could be associated with the inability to generate reactive metabolites.

7.2.5 Proteomic Profiling of Cytotoxicity: Diphenhydramine

For validation purposes, a drug with no-DILI concern was included for comparison using the same cell line and culture techniques. This served to assess whether whole proteome analysis provided too complex a sample and identified significant differences not associated with toxicity. The results in this chapter illustrate augmentation of the HepG2 proteome based solely on culturing technique. Despite this highly advantageous adaption, neither monolayers nor spheroids were capable of identifying proteomic signatures of cytotoxicity related to the three distinct drugs. As the three hepatotoxic drugs were unable to display toxicity, the ability to use Diphenhydramine for validation was limited.

Volcano plots (Figure 7.21) for monolayer cultures, here and for other drugs, were mostly vertical while those for spheroids were slight spread across the x-axis suggesting greater changes in the mean difference. In addition, the protein profiles (Figure 7.21D) for Diphenhydramine were similar to the other drugs. What is apparent is that while validation of model systems is a necessity, here it appears unable to fulfil its purpose.

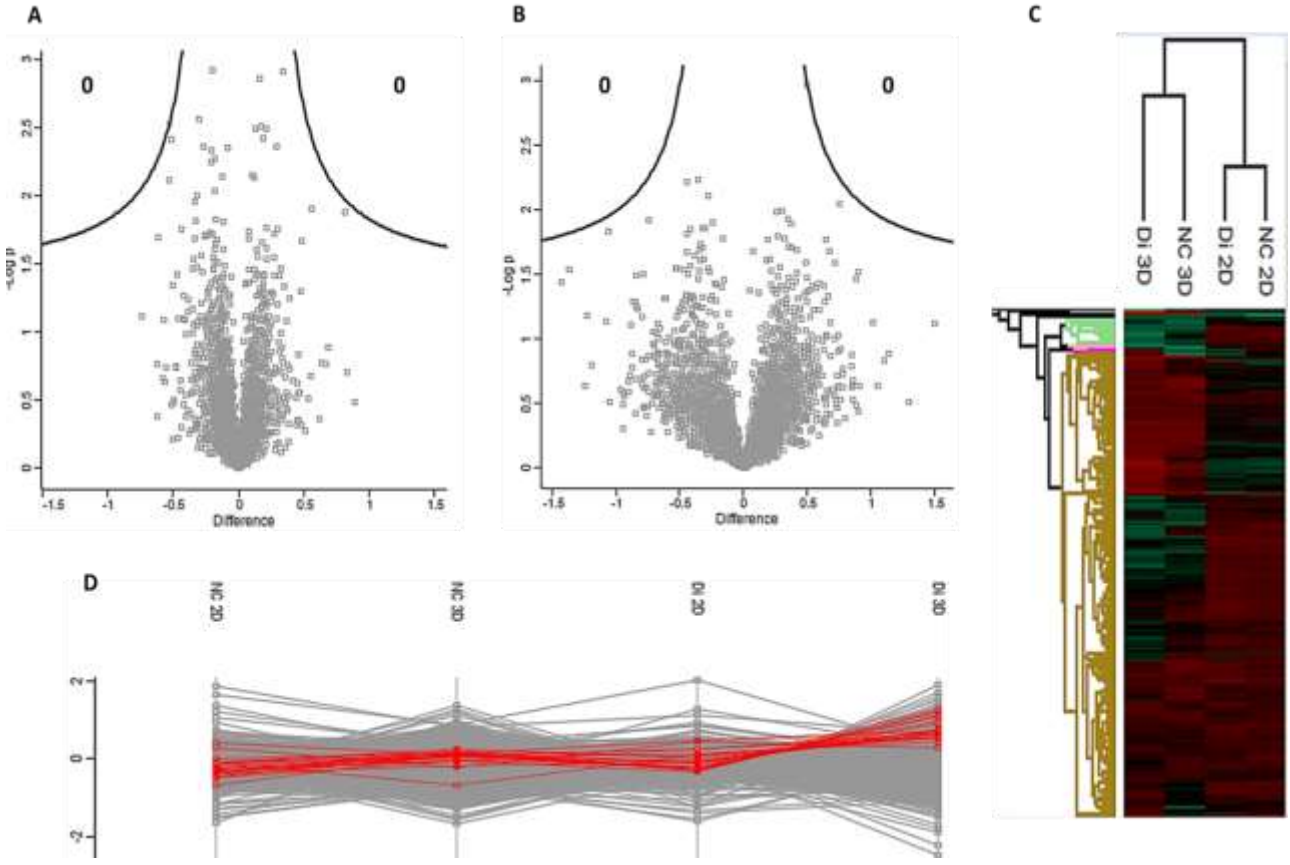

Figure 7.21: Volcano plot comparing HepG2 cells using an FDR of 0.01 with contributions from the p-value and mean difference. A) Negative control monolayers versus Diphenhydramine in monolayers, B) Negative control spheroids versus Diphenhydramine spheroids, C) Hierarchical clustering and D) Protein profiles used to identify proteins increasing in monolayer and cultures exposed to Diphenhydramine.

7.2.6 Overall Implications of Proteomic Profiling of Cytotoxicity

When doing hierarchical clustering, clusters for Bromfenac and Troglitazone (CBT) had the same amount of smaller clusters, containing 0 to 100 proteins, as those for Diphenhydramine and Ketoconazole (CDK). The largest proteins clusters were 2 965, 3 515, 2 820 and 3 404 for Bromfenac, Ketoconazole, Troglitazone and Diphenhydramine respectively. Large clusters with stable expression throughout the protein profile would be advantageous if smaller clusters were able to distinguish hepatotoxicity makers. What was evidenced throughout this chapter is that the samples which were labelled together in the same mass spectrometry experiment demonstrated some acquisition dependent features. To gain a greater overview of the data, PCA (Figure 7.22 and 7.23) was done for all samples analysed from the same labelled replicates.

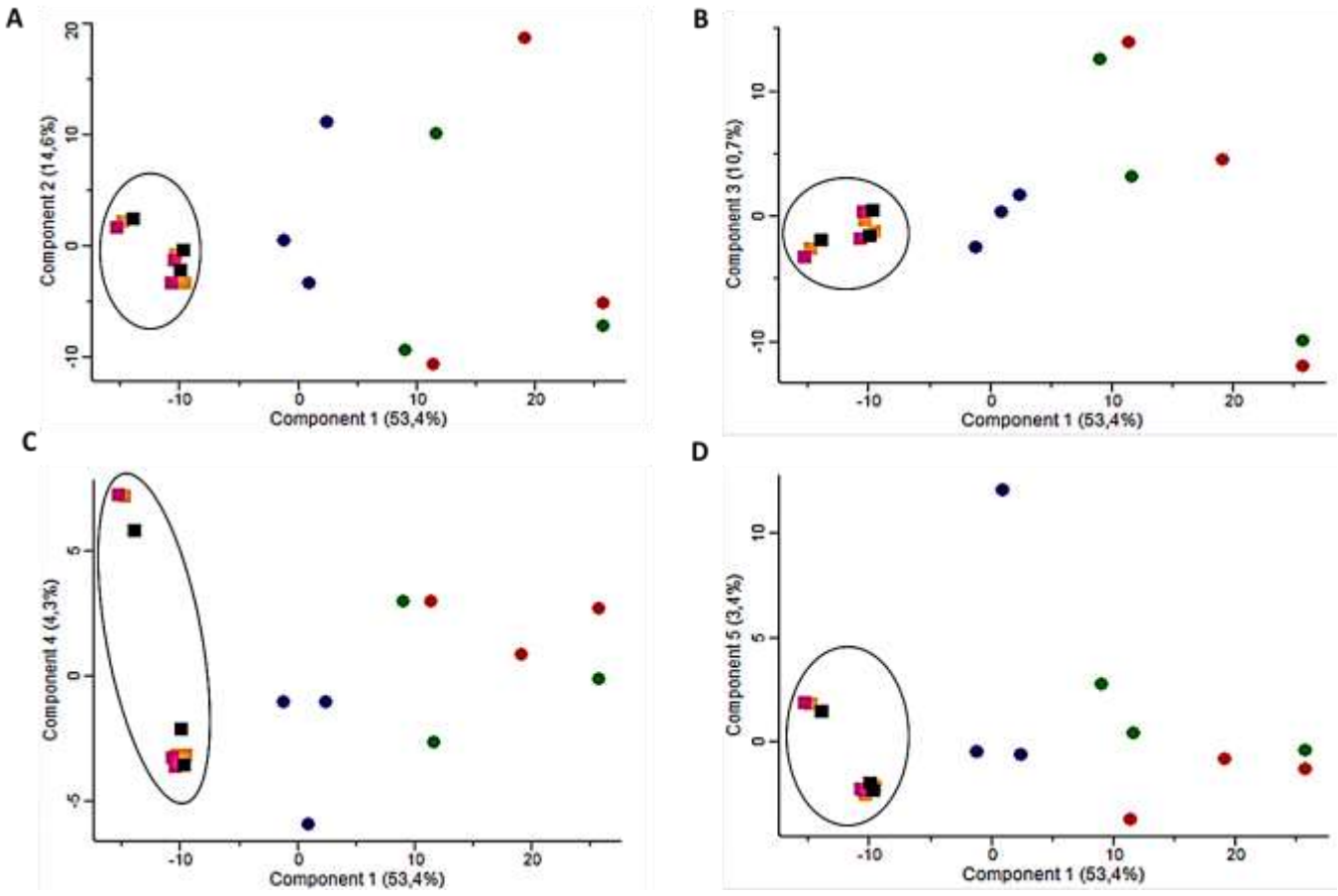

Figure 7.22: Principal component analysis of controls, Bromfenac and Troglitazone (n = 3). A) Component 1 versus Component 2, B) Component 1 versus Component 3, C) Component 1 versus Component 4 and D) Component 1 versus Component 5. Negative control monolayers: black square, Negative control spheroids: red circle, Bromfenac monolayers: orange square, Bromfenac spheroids: green circle, Troglitazone monolayers: pink square and Troglitazone spheroids: blue circle. Components 6 - 17 accounted for 2.3 - 0.6% of the variance.

PCA of monolayers, whether the control or exposed to any of the study drugs, clustered well in the higher principal components. While spheroid samples did not cluster with monolayers there was no tight clustering of spheroids whether exposed to drugs or not. This could highlight a potential inconsistency in the development and maturation of spheroids using the HepG2 cell line.

Despite this largely unpredictable clustering of spheroids, which could be related to a fluctuating HepG2 phenotype, spatially resolving these culturing techniques from one another is important. PCA further reiterated the lack of distinct proteomic profiles generated following exposure to hepatotoxic drugs. This could be due to the large number of proteins masking minor deviations, inadequacies of the model, quantitative noise or a combination of these factors.

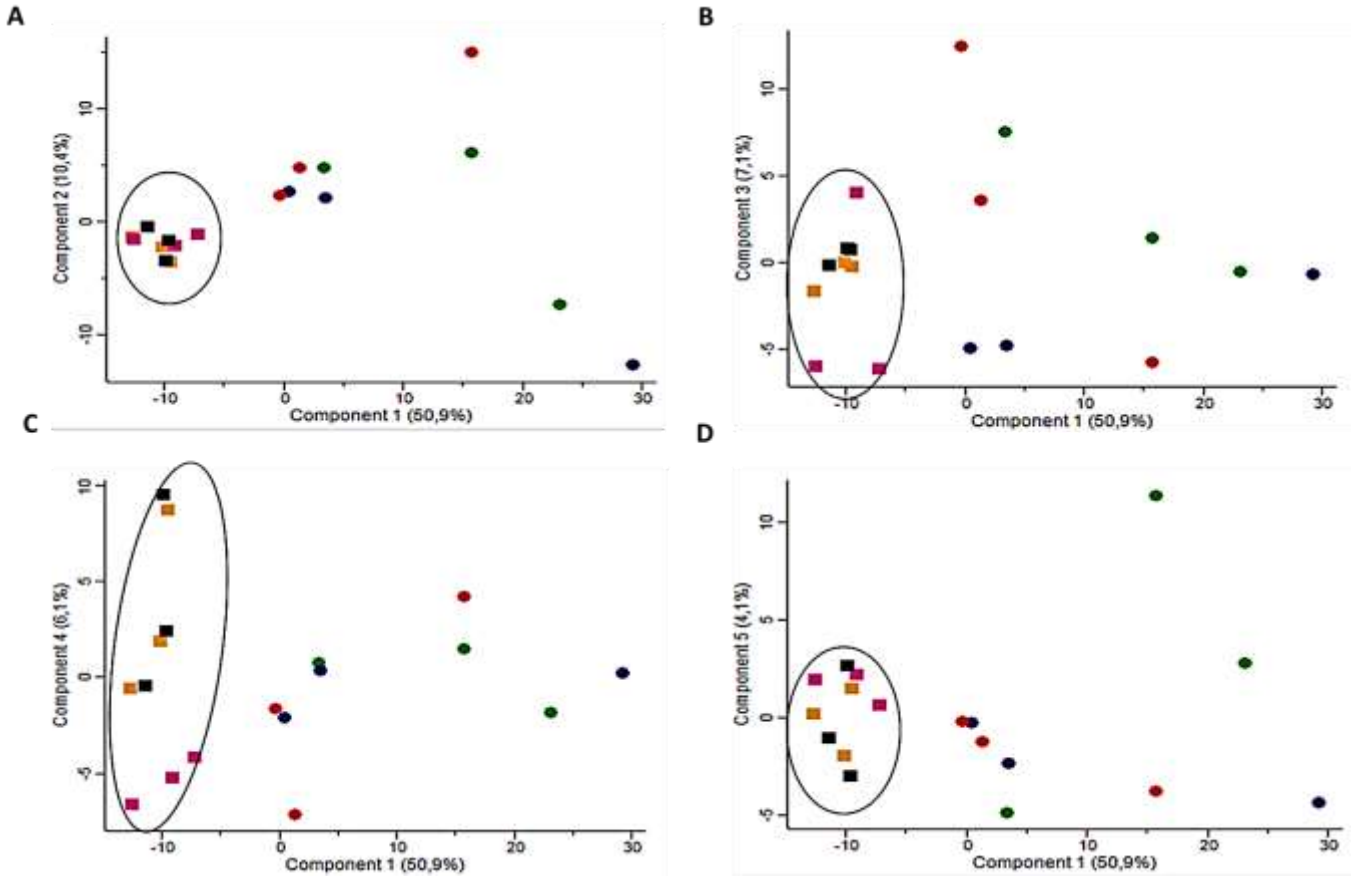

Figure 7.23: Principal component analysis of controls, Diphenhydramine and Ketoconazole (n = 3). A) Component 1 versus Component 2, B) Component 1 versus Component 3, C) Component 1 versus Component 4 and D) Component 1 versus Component 5. Negative control monolayers: black square, Negative control spheroids: red circle, Diphenhydramine monolayers: orange square, Diphenhydramine spheroids: green circle, Ketoconazole monolayers: pink square, Ketoconazole spheroids: blue circle. Components 6 - 17 accounted for 4.0 - 0.5% of the variance.

7.3 Conclusion

Labelled proteomics was successfully used to identify and quantify over 4 600 proteins in a total of six datasets. Over 3 700 proteins were then identified across triplicate experiments using at least 2 unique peptides at an identification confidence of 100%. Culturing HepG2 cells as both monolayers and spheroids demonstrated augmentation of the proteome with approximately 34% of proteins (1 163 of 3 241) being differentially expressed. This supports the fundamental notion that spheroid culture alters vital components of the expressed hepatic phenotype. These changes could be highly dependent on the type of spheroid culture technique and the stability of these cultures over the time course of the experiment. However, the vast extent of this change, which encompasses various protein functions from structural to signalling, demonstrates the ease with which the proteome can be altered based on the cell culture technique and timing of initiation of these cultures.

The notion of spheroid zonation could result in the generation of various proteomes within each spheroid. This potential for multi-proteome expression could suggest that the analysis of spheroids is merely averaging multiple versions of the HepG2 proteome. Various factors could be postulated to contribute to the inability of this model system to provide insight into the proteomic differences expected during hepatotoxic insult using different drug classes known to elicit hepatotoxicity *in vivo*.

Firstly, it has to be considered whether or not this model system is complex enough to detect hepatotoxicity. Based on the findings that very few proteins essential to the three phases of drug metabolism were detected in the HepG2 cell lines assessed, it appears that this model system is not representative of the *in vivo* hepatic proteome, nor complex enough to fulfil the role of identifying proteomic signatures of toxicity where these metabolising enzymes play a critical role in activating the drug through formation of reactive metabolites.

Secondly, were some pathways or mechanisms being overshadowed or missed as a result of comparing such a vast cohort of proteins. In assessing large datasets of thousands of proteins, and many more thousands of peptides, the proteins of lower abundance, although changing significantly may be completely overwhelmed by the more abundant proteins. Potential elucidation of biological variation due to different culturing techniques, as opposed to actual signatures of hepatotoxic protein changes, could also limit the ability to identify the key changes in the proteins showing hepatotoxicity related changes.

Lastly, could short term sub-toxic drug doses be sufficient to trigger the changes in the proteome that provide a signature of hepatotoxicity to unravel the complexities of drug-induced liver injury.

All of this data suggests that, culturing HepG2 cells as static, heterogeneous spheroids is capable of altering the proteome. However, these alterations of the hepatic phenotype, when exposed to DILI inducing drugs after 4 days does not lend itself to the identification of significantly altered proteins indicative of the drugs mechanism of action. This could be due to the fact that spheroids require more time for maturation prior to exposure or that the complexity of this system is insufficient as monolayers or spheroids to provide further insight into drug-induced hepatotoxicity. In addition to these factors, the drugs selected here may be too idiosyncratic in nature to currently model *in vitro.*

Chapter 8: Global Discussion

8.1 Overview

The rapidly changing scientific and regulatory climate has been influenced by two key events in the past decade. Namely, Amendment VII to the European Cosmetics Directive which underwent final implementation in March 2013, and the National Academy of Sciences report entitled 'Toxicity Testing in the 21st Century: A Vision and a Strategy'. As a result, scientists were forced to reassess the use of human cell-based tissue models to reduce and eventually replace the use of animals for toxicity testing.

Combining this with advances in analytical and computational methods resulted in the reinforcement of highly predictive *in vitro* platforms capable of informing regulatory decisions. Innovative cell and tissue models aim to improve extrapolation from *in vitro* human liver models to *in vivo* hepatotoxicity (McKim JM, 2015). However, the paradigm shift to animal replacement, while greatly improved by various models and technologies is yet to prove capable of standing alone in safety predictions.

This study set out to determine the influence of multiple features associated with this paradigm shift in safety predictions. Firstly, could quantitative proteomics methods that employ *in vitro* stable isotope labelling be used to achieve high proteome coverage and track the phenotypic changes during hepatic differentiation from iPSCs. Secondly, could differential culturing of commonly used human hepatocyte cell lines beneficially alter the hepatic proteome.

Thirdly, could quantitative proteomics involving *in vitro* stable isotope labelling be used to achieve high proteome coverage of primary human hepatocytes and differentially sourced human hepatocyte-derived cells to provide biological insights for pre-clinical hepatotoxicity models. Lastly, could this approach be used to achieve high proteome coverage and distinguish drug-specific alteration of the proteome when differentially cultured. This chapter serves to discuss insights influencing study design, perceived shortcomings and future directives.

8.1.1 Mass Spectrometry Quantitation and Labelled Proteomics

Variance in quantitative data can be strongly associated with the sample biology, technical handling and instrument bias. Prior to considering the instrumental and label associated limitations of labelled proteomics, the variance introduced prior to sample analysis needs to be acknowledged. Variance influencing identification and quantitation can be introduced in each step of experimental protocols.

Sample collection and lysis can be the point of origin of potential inconsistency. In a whole cell lysate, highly lipophilic and highly membrane bound proteins may be discarded due to solubility limitations. Protein quantitation using the BCA assay has been found to demonstrate variance of 15 - 30% (Burkhart JM, Vaudel M *et al.*, 2011). Fifty micrograms of protein was labelled per sample. The variance associated with BCA could have resulted in 35 - 65 µg of protein per individual sample.

These differences could be amplified throughout sample preparation, most notably in steps before combining labelled peptides, and have a direct influence on protein quantitation. Protein precipitation can result in intangible loss of specific subsets of proteins, trypsin digestion can be biased with resistance of tightly folded or strong lipid binding proteins to trypsin (missed cleavage ratio lysine:arginine 5:1-6:1) and labelling efficiency can be reduced due to tag hydrolysis.

Following successful labelling, SPE clean-up procedures and fractionation can again reduce the protein cohort due to the various interactions of column chemistries, mobile phases and protein characteristics. Peptide abundance, ionizability, potential for chemical modification, post translational modifications and dynamic range can strongly influence the number of proteins identified. Numerous features associated with variation can be assessed to confirm the validity and integrity of samples including; assessing labelling efficiency, detecting porcine trypsin in samples providing inference for digestion efficiency and observing quantitative variation of housekeeping proteins.

Isobaric tagging for proteomics has become universally acceptable with multiplexing allowing direct ratiometric comparison of independently collected samples. The stochastic nature of data-dependent acquisition, whereby ion intensity dictates selection for fragmentation, presents a challenge with replicates. Systematic variation in experimentation can alter biological conclusions which necessitates appropriate data normalization and analysis (Viner R, Scigelova M *et al.*).

Quantification of complex multiplexed peptide mixtures can be challenging for two major reasons. Residual chemicals can produce background in the low m/z range which interfere with reporter ions from several peptide precursors and potentiate ratio distortion. Secondly, parent ion co-isolation of reporter ions from several peptide precursors also potentiate ratio distortion. These limitations can generally be overcome with high-field Orbitrap mass analysers, advanced signal processing and improved ion optics of newer instruments (Viner R, Scigelova M *et al.*).

Biological insight is dependent on instrument capabilities. Orbitraps, which use an electrostatic field with complex spiral ion movements and oscillation frequencies, surpass the resolution of orthogonal time-of-flight analysers and possess a higher resolving power. Resolving power decreases as m/z increases. However, a resolution of 60 000 can be acquired in under 1 second. Due to the 1 Da mass difference in 6-plex isobaric tag, resolving power of the mass spectrometer is critical (Scigelova M and Makarov A, 2006).

In addition, interfering signals can be reduced or negated by the use of additional gas-phase manipulation such as MS/MS/MS (MS3). The use of isolation wave forms with multiple frequency notches such as synchronous precursor selection (SPS) allows co-isolation and co-fragmented of multiple MS2 fragment ions which ultimately results in improved dynamic range of reporter ion quantitation (McAlister GC, Nusinow DP *et al.*, 2014).

Precision and accuracy are directly influenced by reporter ion liberation. Higher-energy collisional dissociation (HCD) acquisition using multiple collision energies (stepped CE) on C-trap accumulated ions can generate peptide fragments and liberate reporter ions for simultaneous analysis. Stepped CE are beneficial for quantification with increased relative abundance, enhanced resolution and lower relative standard deviation (Zhang T, Jiang L *et al.*, 2015).

Acknowledging these parameters and availability of instrumentation, mass spectrometry was done on a Thermo Scientific QE Hybrid Quadrupole-Orbitrap Mass Spectrometer (Chapter 4: Characterisation of Proteome Changes during Hepatocyte Differentiation and Chapter 5: Proteomic Comparison of Hepatocytes) and a Thermo Scientific Fusion Mass Spectrometer (Chapter 7: Proteomic Profiling of Cytotoxicity).

The Fusion MS allows for faster acquisition speed which logically translates into more PSMs per unit time and theoretically greater protein identification. Despite maintaining parameters as similar as possible between instrumentation, the QE provided more protein identifications (+1500 - 2000) once being processed using identical software. An important feature to note is that for experiments run on the QE hybrid instrument, fifteen sub-fractions of previously chromatographed peptides were analysed using 2 hour elution gradients. While on the Fusion, twelve sub-fractions of previously chromatographed peptides using the same original chromatographic separation but larger fractions were recombined and were analysed using a similar 2 hour elution gradient. Greater sample complexity potentially reduced protein identifications due to a total of 6 less hours of acquisition.

Maximizing proteome coverage requires a balance of replicates versus mass spectrometer acquisition time. Generally, the positive impact on proteome coverage of a second acquisition of fractions in repeat analysis is evidenced (Wang H, Chang-Wong T *et al.*, 2010). Peptide and protein level data increases by approximately 25 - 30% and 11 - 15% respectively with analysis of two replicates with a third technical replicate providing a 5% increase in protein identification (Paulo JA, 2013). While technical repeats positively impact proteome coverage, financial and time implications of experimentation limited technical replicates and only biological replicates were analysed.

Instrument settings can also influence the peptides that are identified and sequences through MS/MS techniques. Due to the complexity of the samples and the high number of peptides that co-elute there is an instrument related bias to the peptides that are analysed and reported, which depends on the individual peptide intensities measured at any given time point during the elution of the peptide mixtures from the column.

Peptides showing higher intensities are preferentially analysed at the expense of the lower abundance peptides eluting at the same time. Instrument settings could inadvertently be chosen such that any low abundance peptides or peptides with poor ionizability or fragmentation characteristics would be ignored by the mass spectrometer control software. Once the precursor ions from co-eluting peptides from the different samples are identified, reporter ions liberated by high collision energy are quantified with protein inference influencing the determination of relative abundance. Unique peptides in a protein can give a direct measure of protein abundance while shared peptides require relatively complex weighted averaging of abundance ratios (Nesvizhskii AI and Aebersold R, 2005).

The influence of analysis platforms and search engines can therefore also not be underestimated. This was demonstrated using the analysis of HLCA replicate 1. Searching 15 MGF files using Proteome Discovered™ Software with Mascot database search algorithms identifying 7 641 proteins. However, using Peptide Shaker with MS-GF+, Comet and X!Tandem only 6 682 proteins were identified despite using the same collected data. This demonstrates that despite a well-designed experiment, biological implications and insight can be skewed or distorted by not implementing the most appropriate bioinformatics strategies.

In addition, missing values and the requirement for normalization complicates labelled proteomics. If a quantitative value is not reported for a specific label it could be due to inefficient labelling, reporter ion intensities being below the lower limit of detection, chemical contaminants masking reporter ions or merely the absence of the protein in the sample.

Missing values, if they are few in number and therefore not associated with inadequate experimental design, can be imputed from a normal distribution but this could alter biological implications if all the inclusions for use are not met. Normalization is required to improve data integrity and ensure comparison between both technical and biological replicates. Most strategies for normalization require various estimations which can be complicated by null intensities. These estimations rely on the competence and integrity of data acquisition and algorithms used to estimate PSM ratios which are used as an inference for peptide and protein ratios. Understanding how these variables may influence and alter results cannot be undervalued.

8.1.2 Manipulating the Hepatic Proteome

The data presented in this study has been captured from several *in vitro* cell cultures under various conditions, including manipulation of differentiation and culturing strategy, which dictate formation and maturation of the hepatic proteome. The complex processes of obtaining a hepatocyte-like cells from undifferentiated precursors was investigated alongside the ability to augment the hepatic proteome of immortalized cell lines. The diverse nature of implementing these two approaches along with the complexities of analysing the proteome of each highlights the necessity for standardization across models for assessing hepatotoxicity.

Complex cellular and non-cellular interactions govern the diverse functionality of the liver which varies over time. Establishing the conditions required to express the fully functional repertoire of hepatocytes requires the understanding of three key factors. These include the multiplicity of cell types, relationships between cellular and non-cellular interactions (ECM) and events modulating gene expression. Co-cultures with non-parenchymal cells promotes both hepatic survival and function, illustrating dependence on cooperative interactions. However, evidence suggests that the ECM may be the most important of these factors (Selden C, Khalil M *et al.*, 1999).

Despite hepatic monolayers being favoured in HTS, the ability to control homotypic and heterotypic interactions in simple spheroid formats may prove highly advantageous. The 3D culture phenotype is complicated by additional considerations of downstream applications and the batch-to-batch variation of many ECM constituents. The need to provide specific constraints that limit heterogeneity necessitates well-defined conditions. While the use of some additives such as reconstituted basement membrane have positively influenced spheroid morphology, their ill-defined nature could produce more variability than is acceptable.

Simplicity in augmenting the hepatic phenotype for proteomic comparison was considered the most beneficial avenue to pursue. This was successfully achieved using HepG2 cells. However, the strength of this altered "*biomimetic microenvironment"* was not sufficient to advance HepG2 cells significantly closer to their differentiated iPSC-HLC counterparts or successors such as PHH.

8.1.3 Liver Tissue Proteome versus Hepatocyte Proteome

Data supports the notion that there is as of yet, no single hepatic model that can accurately reflect the diverse array of outcomes required to mimic the *in vivo* liver. Quantitative label-free shotgun proteomics has compared freshly isolated human hepatocytes with liver tissue where ER membrane proteins including many of the CYP450 enzymes were retained (Vildhede A, Wisniewski JR *et al.*, 2015).

Plasma membrane proteins such as hepatocyte-specific drug transporters and mitochondrial membrane proteins were reduced in isolated hepatocytes compared to liver tissue. Exposure to oxidative stress during isolation procedures are implicated in these differences (Vildhede A, Wisniewski JR *et al.*, 2015). Comparison to PHH as a "*gold standard*" may continue to provide suboptimal imitation of *in vivo* tissue purely based on isolation-induced phenotypic shifts. Therefore, are we striving to mimic a "*gold standard*" that isn't truly gold?

8.1.4 Hepatotoxicity Profiling

Despite not relying on a few molecular or functional endpoints and instead assessing the whole proteome, hepatotoxicity profiling using labelled proteomics did not implicate subsets of proteins in a drug specific manner. Culturing HepG2 cells as both monolayers and spheroids demonstrated alteration in the proteome with approximately 34% of proteins being variably expressed. This supports the fundamental notion that spheroid culture techniques alter the hepatic phenotype. However, these alterations may not be sufficient to draw parallels to functional responses of PHH. Alternatively, the drugs selected for this study were merely too idiosyncratic and combined with the established metabolic incompetence of the cells simply cannot provide a reliable *in vitro* hepatotoxicity model.

A notable absence of phase I metabolizing proteins was observed in HepG2 cells cultured as monolayers or spheroids (Chapter 7). As these proteins are generally lipophilic and membrane-associated it was considered possible that potential bias due to the fractionation and recombination used in sample preparation for LC-MS/MS may have resulted in a loss of this subset of proteins. However, when comparing the HepG2 data from the hepatocyte comparison (Chapter 5) to the proteomic profiling of cytotoxicity (Chapter 7), it was noted that the CYP proteins with the highest abundance Chapter 5 were the ones identified in Chapter 7.

This suggests that the cohort of metabolising enzymes are reduced in abundance and expression with whole proteome analysis potentially masking detection. Combining a targeted mass spectrometry approach with the instrument auto-selection used in this study could aid in the identification and quantification of these important metabolizing enzymes.

Bhatia *et al.* illustrated that proximity to heterotypic interfaces promotes hepatic functions of PHHs in micropatterned co-cultures (MPCCs)(Bhatia S, Balis U *et al.*, 1999). Stable human MPCCs (7 days), with repeated dosing, have demonstrated a 65.7% sensitivity in detecting DILI with an FDR of 10%. Despite only detecting 67.5% of the drugs inducing DILI, this method shows significant improvement in sensitivity predictions (Khetani SR, Kanchagar C *et al.*, 2013). With PHH MPCCs detecting only 67.5% of DILI, it is possible that the system used in this study is merely not complex or sensitive enough to achieve detectable hepatotoxic effects within the proteome.

Understandably, monolayer cultures did not elucidate hepatotoxic features. Despite altering the expression of 1 100+ proteins by culturing HepG2 cells as spheroids this did not aid in predicating hepatotoxicity. MPCCs (500 µm diameter, 1200 µm centre-to-centre spacing) of iPSC-HLCs with hepatic polygonal morphology, are reported to display varying degrees of phase I and II metabolising enzymes and down-regulation of foetal markers. When used by others in DILI predictions, HLC micropatterned co-cultures reveal similarity to PHH with 65 - 70% sensitivity and 100% specificity. This demonstrates that HLCs, if cultured appropriately, could be valuable for drug discovery applications (Ware BR, Berger DR *et al.*, 2015). The ability to augment the proteome of HepG2 cells and the increased sensitivity and specificity of HLCs suggests that growth and characterisation of HLCs as spheroids will be advantageous in mimicking PHHs.

Idiosyncratic hepatotoxicity is rare (<0.1%) and typically not associated with the known pharmacological properties of the drug. While ill-defined, distinct steps in generating idiosyncratic drug reactions include target tissue exposure, reactive intermediate formation, inadequate detoxification, covalent binding to macromolecules, formation and immune recognition of neoantigens, formation of cytotoxic antibodies and subsequent liver toxicity (Walgren JL, Mitchell MD *et al.*, 2005). This diverse array of factors has proven impossible to mimic or predict in a single preclinical model. Results from the drugs used in this study pose various interesting questions. Is our model complex enough to detect DILI? Does whole proteome analysis mask potential protein signatures? Or are these drugs just not detected as hepatotoxic because it is impossible to mimic idiosyncratic phenomena *in vitro*?

8.2 Conclusions

To the extent of our knowledge, there is no published data which has as comprehensively detailed the hepatic proteome of hepatocytes from various origins with noteworthy replicates. In addition, this is one of the first studies to systematically examine proteomic differentiation from induced pluripotent stem cells into hepatic lineages. Objectives undertaken, and achieved in this project, included; successful adoption and characterisation of a 3D culture model of HepG2 cells, differentiation of iPSCs into hepatocyte-like cells with complementary proteomic data tracking phases of differentiation, multi-parametric assaying of hepatotoxins and labelled proteomics to capture expansive whole proteome data to answer various biological questions associated with hepatotoxicity models.

This research was divided into two distinct aims. Firstly, to determine proteomic differences between primary human hepatocytes and differentially cultured and sourced human hepatocyte-derived cell lines or differentiated hepatocyte-like cells. This included characterisation of iPSC-HLCs over the differentiation time courses and included several biological repeats. These assessments revealed a potential limited proteomic window for maintenance of the hepatic phenotype in HLCs. This window may not be correlated to RNA data, or be inclusive of all proteins and be biased by specific differentiations, but does create a timeframe of applicability.

Globally, HLCs appeared to mimic PHH better than the HepG2 cell line models despite comparing with cells collected beyond the proteomic window suggested. HepG2 spheroids, unlike HepG2 monolayers, were capable of displaying liver specific functions similar to HLCs. Ranking cells in order of equivalence to PHHs may rank HLCs and HepG2 spheroids similarly but above HepG2 monolayers. Repeating this experiment with differentiation of HLCs terminated a few days earlier may see HLCs outperforming HepG2 spheroids.

The second aim was to use a cellular hepatocyte model (HepG2) to generate non-specific, proteome-wide information associated with exposure to selected hepatotoxins to identify potential proteomic signatures of hepatotoxicity. While culturing in 3D demonstrated the feasibility of altering the proteome based on culture method, the inability of this model system to provide insight into the proteomic changes that take place during hepatotoxicity can be associated with various factors. In summation, HepG2 spheroids may not be complex enough to uncover hepatotoxicity or the mechanism of action of the study drugs are too multifaceted for current in *vitro* model systems to identify.

As of yet there appears to be no definitive solution to modelling proteomic signatures of hepatotoxicity. However, collection of quantitative whole proteome data means that the full depth of biological insight is yet to be uncovered. The ability to generate iPSC-HLCs with an established phenotype and for 3D cultures to increase correlation of immortalized cell lines to primary hepatocytes suggests that the combination of these factors may be ideal in capturing the hepatic proteome closest to that of PHH.

8.3 Study Shortcomings

Constraints in scientific experimentation can be built-in by poor study design and ultimately result in the inability to extrapolate the model to the *in vivo* scenario. Despite this project being designed to minimize obvious shortcomings, logistics and feasibility resulted in some concerns which can be seen as limiting. These are discussed below and are highlighted under three distinct areas being cells and cell culture, proteomics and hepatotoxicity profiling.

8.3.1 Cells and Cell Culture

Throughout the study, two HepG2 cell lines were used which may have created inconsistency in the results. One HepG2 cell line, purchased from the European Collection of Authenticated Cell Cultures (ECACC; Wiltshire, UK) was used for the proteomic comparisons between cell types (PHHs, HLCs and HepG2 cells). While a second HepG2 cell line purchased from Cellonex (South Africa) was used for proteomic profiling during the hepatotoxicity testing. Despite being an immortalized cell line which is required to meet certain criteria, the varied origin could be a notable shortcoming. Attempted maturation of the hepatic phenotype by 3D culture was limited to approximately 14 days due to the static nature of this culture method. Extensive maturation has been achieved over a longer time course in more complex dynamic systems with enhanced nutrient, waste and oxygen exchange.

During this study it was not possible to culture and characterise HLCs as spheroids. It has been shown that culturing HLCs as spheroids promotes a more mature adult phenotype. However, the degree of maturation could be highly dependent on the culture technique. Therefore, features associated with hepatocyte maturation could not be inferred from another method. This limited the use of HLC spheroids which were produced in hanging drops similarly to HepG2 spheroids but not characterised. An additional limitation of HLCs was the collection time point for comparison with other cells. If the proteomics time course had been analysed prior to the hepatocyte comparison samples would have been collected at approximately Day 32 as opposed to Day 35 at which time the phenotype appeared to be moving away from the typical hepatocyte phenotype seen between Day 28 and Day 32.

8.3.2 Proteomics

Proteomic experiments using isobaric labelling *in vitro* are associated with great cost and long data acquisition time. This creates a problem in assessing the importance of biological and technical replicates. The extent of comparisons conducted (thirteen 6-plex TMT labelled samples including biological replicates with 354 hours of MS acquisition time) in this study limited the time to perform technical repeats. The iPSC-HLC time-courses, while designed to have an overlap of duplicates for some tags, only had a single quantitation value for some time points throughout differentiation. Despite this, data trends between time courses correlated well, however, these relative quantitative results need to be considered cautiously.

The observations of hepatic markers during the time courses should have dictated the collection of terminally-differentiated HLCs. However, as these experiments were conducted concurrently, Day 35 HLCs were compared to PHHs and HepG2 cells. This could underestimate the proteomic correlation of iPSC-HLCs to PHHs. Proteomics was conducted across various laboratories which influenced fractionation of TMT-labelled samples, sample pooling strategies and LC-MS/MS analysis. Samples for the iPSCs-HLC time-courses and hepatocyte comparisons were fractioned using a UPLC system, pooled into 15 fractions across the gradient to achieve well distributed and approximately equal peptide concentrations across the elution gradient. These recombined sub-fractions were analysed on a Q Exactive Orbitrap mass spectrometer system.

Samples for hepatotoxicity profiling were fractioned using an HPLC system, and in the same manner were recombined but to give only 12 sub-fractions. The pooled sub-fractions were analysed on a Fusion Orbitrap mass spectrometer system. The enhanced complexity per sub-fraction for hepatotoxicity profiling due to having 3 less fractions (6 less hours of MS acquisition time) influenced the number of proteins identified. Detection of low abundance proteins would have been reduced based on this and could have significant implications in the identification and quantitation of low abundance peptides. Based on this, targeted proteomics using these samples can still be done to elucidate pathways of interest hypothesised to have an influence during the hepatotoxicity testing.

8.3.3 Hepatotoxicity Profiling

Validation is an important feature of designing *in vitro* models. This is why a non-hepatotoxic drug was selected to determine whether potential alterations by hepatotoxic drugs were a result of experimental methods. While in theory Diphenhydramine was capable of providing this validation, the use of toxic/non-toxic pair from the same class of drugs such as Troglitazone and Rosiglitazone could prove to be more insightful in a model capable of detecting hepatotoxicity.

Having only cultured cells as spheroids for 4 days prior to exposure could have been insufficient to adopt adequate features of hepatocyte maturation resulting in the inability to detect proteomic differences associated with each drug in spheroid cultures. However, it remains possible that global differences seen in the 2D versus 3D phenotypes were too vast and that the altered molecular pathways involved in early hepatotoxicity are masked by the more abundant changes due to the culturing techniques used.

In addition to the biological features of hepatotoxicity profiling, samples run on the Fusion MS could have been analysed using SPS-MS3. Making use of an additional fragmentation improves the dynamic range of reporter ion quantitation and reduces reporter ion signal variance. This diminishes quantitative noise and potentially could have improved data accuracy and enhanced the informational value of profiling experiments.

8.4 Future Directives

Research currently underway includes looking at enhancing spheroid compaction and aggregation using various 3D culture formats with and without addition of natural or synthesis ECM. Factors such as seeding density and culture duration appear to have as strong an influence on cultures as the addition of ECM which influences the gross morphology of spheroids. Furthermore, research is being conducted to observe whether using these cell culture methods with well-defined prototypic inducers can alter or enhance the cohort of metabolising enzymes at the proteomic and metabolomic level.

Determining which cells most reflect PHHs is merely the initial step in defining valid hepatotoxicity models. The liver is composed of approximately one-third non-parenchymal cells and two-thirds hepatocytes and as such the core of liver formation, function, physiology and pathophysiology is heterotypic interactions. Proximity to heterotypic interfaces has been shown to promote hepatic functions suggesting that an appropriate co-culture would be advantageous to the hepatic phenotype. Therefore, future research will be based on comparing liver tissue to hepatocytes with various combinations of other liver resident cells.

Additionally, if efficiently characterized, mature HLCs can be grown using hanging drops or similar 3D culture platforms for use in co-cultures. As with other hepatocyte types, HLC's do not proliferate in culture and therefore could be configured with growth arrested non-parenchymal cells to further enhance the ability to mimic liver tissue by the presence of heterotypic configurations.

References

3D-Biomatrix (2013). 3D Cell Culture 101: An introduction to 3D cell culture tools and techniques.

Abboud G and Kaplowitz N (2007). "Drug-induced liver injury." *Drug Safety* **30**(4): 277-94.

Achour B, Russell MR, Barber J and Rostami-Hodjegan A (2014). "Simultaneous quantification of the abundance of several cytochrome P450 and uridine 5'-diphospho-glucuronosyltransferase enzymes in human liver microsomes using multiplexed targeted proteomics." *Drug Metabolism and Disposition* **42**(4): 500-10.

Agarwal S, Holton KL and Lanza R (2008). "Efficient differentiation of functional hepatocytes from human embryonic stem cells." *Stem Cells* **26**(5): 1117-27.

Ahuja V and Sharma S (2014). "Drug safety testing paradigm, current progress and future challenges: an overview." *Journal of Applied Toxicology* **34**(6): 576-94.

Altelaar A and Heck AJ (2012). "Trends in ultrasensitive proteomics." *Current Opinion in Chemical Biology* **16**(1): 206-13.

Anson BD, Kolaja K and Kamp TJ (2011). "Opportunities for human iPS cells in predictive toxicology." *Clinical Pharmacology and Therapeutics* **89**(5): 754-58.

Au J, Navarro V and Rossi S (2011). "Review article: Drug-induced liver injury – its pathophysiology and evolving diagnostic tools." *Alimentary Pharmacology and Therapeutics* **34**(1): 11-20.

Ballet F (1997). "Hepatotoxicity in drug development: Detection, significance and solutions." *Journal of Hepatology* **26**: 26-36.

Baxter M, Withey S, Harrison S, Segeritz CP, Zhang F, Atkinson-Dell R, Rowe C, Gerrard DT, Sison-Young R and Jenkins R (2015). "Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes." *Journal of Hepatology* **62**(3): 581-89.

Benevento M and Munoz J (2012). "Role of mass spectrometry-based proteomics in the study of cellular reprogramming and induced pluripotent stem cells." *Expert Review of Proteomics* **9**(4): 379-99.

Bhadriraju K and Chen CS (2002). "Engineering cellular microenvironments to improve cell-based drug testing." *Drug Discovery Today* **7**(11): 612-20.

Bhatia S, Balis U, Yarmush M and Toner M (1999). "Effect of cell–cell interactions in preservation of cellular phenotype: Cocultivation of hepatocytes and nonparenchymal cells." *The FASEB Journal* **13**(14): 1883-900.

Burkhart JM, Vaudel M, Zahedi RP, Martens L and Sickmann A (2011). "iTRAQ protein quantification: A quality-controlled workflow." *Proteomics* **11**(6): 1125-34.

Cai J, DeLaForest A, Fisher J, Urick A, Wagner T, Twaroski K, Cayo M, Nagaoka M and Duncan SA (2012). "Protocol for directed differentiation of human pluripotent stem cells toward a hepatocyte fate." *StemBook*.

Carrillo B, Yanofsky C, Laboissiere S, Nadon R and Kearney RE (2010). "Methods for combining peptide intensities to estimate relative protein abundance." *Bioinformatics* **26**(1): 98-103.

Cavnar SP, Salomonsson E, Luker KE, Luker GD and Takayama S (2013). "Transfer, imaging, and analysis plate for facile handling of 384 hanging drop 3D tissue spheroids." *Journal of Laboratory Automation* **1**: 1-7.

Chadeau-Hyam M, Campanella G, Jombart T, Bottolo L, Portengen L, Vineis P, Liquet B and Vermeulen RC (2013). "Deciphering the complex: Methodological overview of statistical models to derive OMICS-based biomarkers." *Environmental and Molecular Mutagenesis* **54**(7): 542-57.

Chang C and Schiano T (2007). "Review article: Drug hepatotoxicity." *Alimentary Pharmacology and Therapeutics* **25**(10): 1135-51.

Chang TT and Hughes-Fulford M (2014). "Molecular mechanisms underlying the enhanced functions of three-dimensional hepatocyte aggregates." *Biomaterials* **35**(7): 2162-71.

Chen M, Vijay V, Shi Q, Liu Z, Fang H and Tong W (2011). "FDA-approved drug labeling for the study of drug-induced liver injury." *Drug Discovery Today* **16**(15): 697-703.

Costa A, Sarmento B and Seabra V (2014). "An evaluation of the latest *in vitro* tools for drug metabolism studies." *Expert Opinion on Drug Metabolism and Toxicology* **10**(1): 103-19.

Dagda RK, Sultana T and Lyons-Weiler J (2010). "Evaluation of the consensus of four peptide identification algorithms for tandem mass spectrometry based proteomics." *Journal of Proteomics and Bioinformatics* **3**(5): 39-47.

Dambach DM, Andrews BA and Moulin F (2005). "New technologies and screening strategies for hepatotoxicity: Use of *in vitro* models." *Toxicologic Pathology* **33**(1): 17-26.

Davila JC, Rodriguez RJ, Melchert RB and Acosta Jr D (1998). "Predictive value of *in vitro* model systems in toxicology." *Annual Review of Pharmacology and Toxicology* **38**(1): 63-96.

DeLaForest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA and Duncan SA (2011). "HNF4-alpha is essential for specification of hepatic progenitors from human pluripotent stem cells." *Development and Stem Cells* **138**(19): 4143-53.

Demartini DR (2013). A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses. Tandem Mass Spectrometry - Molecular Characterization.

Dhaliwal A (2012). "3D Cell Culture: A Review." *Labome: Mater Methods* **2**(162).

DiMasi JA, Hansen RW and Grabowski HG (2003). "The price of innovation: New estimates of drug development costs." *Journal of Health Economics* **22**(2): 151-85.

Duret C, Gerbal-Chaloin S, Ramos J, Fabre JM, Jacquet E, Navarro F, Blanc P, Sa-Cunha A, Maurel P and Daujat-Chavanieu M (2007). "Isolation, characterization, and differentiation to hepatocyte-like cells of nonparenchymal epithelial cells from adult human liver." *Stem Cells* **25**(7): 1779-90.

Feldmann G (1997). "Liver apoptosis." *Journal of Hepatology* **26**: 1-11.

Feng Y, De Franceschi G, Kahraman A, Soste M, Melnik A, Boersema PJ, de Laureto PP, Nikolaev Y, Oliveira AP and Picotti P (2014). "Global analysis of protein structural changes in complex proteomes." *Nature Biotechnology* **32**(10): 1036-44.

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

Fontana RJ, McCashland TM, Benner KG, Appelman HD, Gunartanam NT, Wisecarver JL, Rabkin JM and Lee WM (1999). "Acute liver failure associated with prolonged use of bromfenac leading to liver transplantation." *Liver Transplantation and Surgery* **5**(6): 480-84.

Fox D, Morris L, Nystul T and Spradling A (2009). "Lineage analysis of stem cells." *StemBook*.

Fukami T, Iida A, Konishi K and Nakajima M (2016). "Human arylacetamide deacetylase hydrolyzes ketoconazole to trigger hepatocellular toxicity." *Biochemical Pharmacology* **116**: 153-61.

Gieseck RL, Hannan NR, Bort R, Hanley NA, Drake RA, Cameron GW, Wynn TA and Vallier L (2014). "Maturation of induced pluripotent stem cell derived hepatocytes by 3D-culture." *PloS One* **9**(1): 1-7.

Gieseck RL, Vallier L and Hannan NR (2015). "Generation of hepatocytes from pluripotent stem cells for drug screening and developmental modeling." *Protocols in In Vitro Hepatocyte Research*: 123-42.

Gómez-Lechón MJ, Tolosa L, Castell JV and Donato MT (2010). "Mechanism-based selection of compounds for the development of innovative *in vitro* approaches to hepatotoxicity studies in the LIINTOP project." *Toxicology in Vitro* **24**(7): 1879-89.

Guillouzo A, Morel F, Langouët S, Maheo K and Rissel M (1997). "Use of hepatocyte cultures for the study of hepatotoxic compounds." *Journal of Hepatology* **26**: 73-80.

Guo L, Dial S, Shi L, Branham W, Liu J, Fang J-L, Green B, Deng H, Kaput J and Ning B (2010). "Similarities and differences in the expression of drug metabolizing enzymes between human hepatic cell lines and primary human hepatocytes." *Drug Metabolism and Disposition* **39**(3): 528-38.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH and Aebersold R (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." *Nature Biotechnology* **17**(10): 994-99.

Hannan NR, Segeritz CP, Touboul T and Vallier L (2013). "Production of hepatocyte-like cells from human pluripotent stem cells." *Nature Protocols* **8**(2): 430-37.

Härmä V, Virtanen J, Mäkelä R, Happonen A, Mpindi J-P, Knuuttila M, Kohonen P, Lötjönen J, Kallioniemi O and Nees M (2010). "A comprehensive panel of three-dimensional models for studies of prostate cancer growth, invasion and drug responses." *PloS One* **5**(5): 1-17.

Hart SN, Li Y, Nakamoto K, Subileau E, Steen D and Zhong X (2010). "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* **38**(6): 988-94.

Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriguen A, Black JR, Elcombe C, Ross JA and Wolf R (2008). "Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development *in vivo*." *Stem Cells* **26**(4): 894-902.

Hirschhaeuser F, Menne H, Dittfeld C, West J, Mueller-Klieser W and Kunz-Schughart LA (2010). "Multicellular tumor spheroids: An underestimated tool is catching up again." *Journal of Biotechnology* **148**(1): 3-15.

Hunter EB, Johnston PE, Tanner G, Pinson CW and Awad JA (1999). "Bromfenac (Duract)-associated hepatic failure requiring liver transplantation." *The American Journal of Gastroenterology* **94**(8): 2299-301.

Ivascu A and Kubbies M (2006). "Rapid generation of single-tumor spheroids for high-throughput cell function and toxicity analysis." *Journal of Biomolecular Screening* **11**(8): 922-32.

Ivascu A and Kubbies M (2007). "Diversity of cell-mediated adhesions in breast cancer spheroids." *International Journal of Oncology* **31**(6): 1403-13.

Jones DP, Lemasters JJ, Han D, Boelsterli UA and Kaplowitz N (2010). "Mechanisms of pathogenesis in drug hepatotoxicity putting the stress on mitochondria." *Molecular Interventions* **10**(2): 98-111.

Kampf C, Mardinoglu A, Fagerberg L, Hallström BM, Edlund K, Lundberg E, Pontén F, Nielsen J and Uhlen M (2014). "The human liver-specific proteome defined by transcriptomics and antibody-based profiling." *The FASEB Journal* **28**(7): 2901-14.

Kassahun K, Pearson PG, Tang W, McIntosh I, Leung K, Elmore C, Dean D, Wang R, Doss G and Baillie TA (2001). "Studies on the metabolism of troglitazone to reactive intermediates *in vitro* and *in vivo*. Evidence for novel biotransformation pathways involving quinone methide formation and thiazolidinedione ring scission." *Chemical Research in Toxicology* **14**(1): 62-70.

Kelly JH and Darlington GJ (1989). "Modulation of the liver specific phenotype in the human hepatoblastoma line HepG2." *In Vitro Cellular and Developmental Biology* **25**(2): 217-22.

Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT, Lorenz K, Lee EH, Barcellos-Hoff MH and Petersen OW (2007). "The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression." *Molecular Oncology* **1**(1): 84-96.

Kessner D, Chambers M, Burke R, Agus D and Mallick P (2008). "ProteoWizard: Open source software for rapid proteomics tools development." *Bioinformatics* **24**(21): 2534-36.

Khetani SR and Bhatia SN (2008). "Microscale culture of human liver cells for drug development." *Nature Biotechnology* **26**(1): 120-26.

Khetani SR, Kanchagar C, Ukairo O, Krzyzewski S, Moore A, Shi J, Aoyama S, Aleo M and Will Y (2013). "The use of micropatterned co-cultures to detect compounds that cause drug induced liver injury in humans." *Toxicological Sciences* **132**(1): 107-17.

Kim JB (2005). "Three-dimensional tissue culture models in cancer biology." *Seminars in Cancer Biology* **15**(5): 365-77.

Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R and Jain S (2014). "A draft map of the human proteome." *Nature* **509**: 575-81.

Kinter LB and Valentin JP (2002). "Safety pharmacology and risk assessment." *Fundamental and Clinical Pharmacology* **16**(3): 175-82.

Knasmüller S, Mersch-Sundermann V, Kevekordes S, Darroudi F, Huber W, Hoelzl C, Bichler J and Majer B (2004). "Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants: Current state of knowledge." *Toxicology* **198**(1): 315-28.

Kola I and Landis J (2004). "Can the pharmaceutical industry reduce attrition rates?" *Nature Reviews Drug Discovery* **3**(8): 711-16.

Larance M and Lamond AI (2015). "Multidimensional proteomics for cell biology." *Nature Reviews Molecular Cell Biology* **16**(5): 269-80.

Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM and Bor DH (2002). "Timing of new black box warnings and withdrawals for prescription medications." *JAMA* **287**(17): 2215-20.

LeCluyse EL, Witek RP, Andersen ME and Powers MJ (2012). "Organotypic liver culture models: Meeting current challenges in toxicity testing." *Critical Reviews in Toxicology* **42**(6): 501-48.

Lee KD, Kuo TKC, Whang-Peng J, Chung YF, Lin CT, Chou SH, Chen JR, Chen YP and Lee OKS (2004). "*In vitro* hepatic differentiation of human mesenchymal stem cells." *Hepatology* **40**(6): 1275-84.

Lemaigre F (2010). "Markers and signaling factors for stem cell differentiation to hepatocytes: Lessons from developmental studies." *Hepatocytes: Methods and Protocols* **640**: 157-66.

Li VC and Kirschner MW (2014). "Molecular ties between the cell cycle and differentiation in embryonic stem cells." *Proceedings of the National Academy of Sciences* **111**(26): 9503-08.

Lin R-Z, Chou LF, Chien CCM and Chang HY (2006). "Dynamic analysis of hepatoma spheroid formation: roles of E-cadherin and β1-integrin." *Cell and Tissue Research* **324**(3): 411-22.

MacCoss MJ and Yates JR (2001). "Proteomics: Analytical tools and techniques." *Current Opinion in Clinical Nutrition and Metabolic Care* **4**(5): 369-75.

Maherali N and Hochedlinger K (2008). "Guidelines and techniques for the generation of induced pluripotent stem cells." *Cell Stem Cell* **3**(6): 595-605.

Masubuchi Y (2006). "Metabolic and non-metabolic factors determining troglitazone hepatotoxicity: a review." *Drug Metabolism and Pharmacokinetics* **21**(5): 347-56.

McAlister GC, Nusinow DP, Jedrychowski MP, Wühr M, Huttlin EL, Erickson BK, Rad R, Haas W and Gygi SP (2014). "MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes." *Analytical Chemistry* **86**(14): 7150-58.

McKim JM (2015). "The journal of applied *in vitro* toxicology: Its time has come!" *Applied In Vitro Toxicology* **1**(1): 1-4.

Mehta G, Hsiao AY, Ingram M, Luker GD and Takayama S (2012). "Opportunities and challenges for use of tumor spheroids as models to test drug delivery and efficacy." *Journal of Controlled Release* **164**(2): 192-204.

Meyer B, Papasotiriou DG and Karas M (2011). "100% protein sequence coverage: A modern form of surrealism in proteomics." *Amino Acids* **41**(2): 291-310.

Mills JB, Rose KA, Sadagopan N, Sahi J and de Morais SM (2004). "Induction of drug metabolism enzymes and MDR1 using a novel human hepatocyte cell line." *Journal of Pharmacology and Experimental Therapeutics* **309**(1): 303-09.

Moore RN, Dasgupta A, Rajaei N, Yarmush ML, Toner M, Larue L and Moghe PV (2008). "Enhanced differentiation of embryonic stem cells using co-cultivation with hepatocytes." *Biotechnology and Bioengineering* **101**(6): 1332-43.

Moses PL, Schroeder B, Alkhatib O, Ferrentino N, Suppan T and Lidofsky SD (1999). "Severe hepatotoxicity associated with bromfenac sodium." *The American Journal of Gastroenterology* **94**(5): 1393-96.

Nesvizhskii AI and Aebersold R (2005). "Interpretation of shotgun proteomic data: The protein inference problem." *Molecular and Cellular Proteomics* **4**(10): 1419-40.

Niles AL, Moravec RA and Riss TL (2009). "*In vitro* viability and cytotoxicity testing and same-well multi-parametric combinations for high throughput screening." *Current chemical genomics* **3**: 33-41.

Olsavsky KM, Page JL, Johnson MC, Zarbl H, Strom SC and Omiecinski CJ (2007). "Gene expression profiling and differentiation assessment in primary human hepatocyte cultures, established hepatoma cell lines, and human liver tissues." *Toxicology and Applied Pharmacology* **222**(1): 42-56.

Olson H, Betton G, Robinson D, Thomas K, Monro A, Kolaja G, Lilly P, Sanders J, Sipes G and Bracken W (2000). "Concordance of the toxicity of pharmaceuticals in humans and in animals." *Regulatory Toxicology and Pharmacology* **32**(1): 56-67.

Park BK, Kitteringham NR, Maggs JL, Pirmohamed M and Williams DP (2005). "The role of metabolic activation in drug-induced hepatotoxicity." *Annual Review of Pharmacology and Toxicology* **45**: 177-202.

Paulo JA (2013). "Practical and efficient searching in proteomics: A cross engine comparison." *WebmedCentral* **4**(10): 1-15.

Peters TS (2005). "Do preclinical testing strategies help predict human hepatotoxic potentials?" *Toxicologic Pathology* **33**(1): 146-54.

Pichler P, Köcher T, Holzmann J, Mazanek M, Taus T, Ammerer G and Mechtler K (2010). "Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap." *Analytical Chemistry* **82**(15): 6549-58.

Piersma SR, Warmoes MO, de Wit M, de Reus I, Knol JC and Jiménez CR (2013). "Whole gel processing procedure for GeLC-MS/MS based proteomics." *Proteome Science* **11**(1): 1-8.

Pottiez G, Wiederin J, Fox HS and Ciborowski P (2012). "Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples." *Journal of Proteome Research* **11**(7): 3774-81.

Qattan AT, Mulvey C, Crawford M, Natale DA and Godovac-Zimmermann J (2010). "Quantitative organelle proteomics of MCF-7 breast cancer cells reveals multiple subcellular locations for proteins in cellular functional processes." *Journal of Proteome Research* **9**(1): 495-508.

Ramaiahgari SC, den Braver MW, Herpers B, Terpstra V, Commandeur JN, van de Water B and Price LS (2014). "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* **88**(5): 1083-95.

Rashid ST, Corbineau S, Hannan N, Marciniak SJ, Miranda E, Alexander G, Huang-Doran I, Griffin J, Ahrlund-Richter L and Skepper J (2010). "Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells." *The Journal of Clinical Investigation* **120**(9): 3127-36.

Rashid ST and Vallier L (2010). "Induced pluripotent stem cells – alchemist's tale or clinical reality?" *Expert Reviews in Molecular Medicine* **12**: 1-14.

Redfern WS, Wakefield ID, Prior H, Pollard CE, Hammond TG and Valentin JP (2002). "Safety pharmacology – a progressive approach." *Fundamental and Clinical Pharmacology* **16**(3): 161-73.

Rodriguez RJ and Acosta D (1997). "Metabolism of ketoconazole and deacetylated ketoconazole by rat hepatic microsomes and flavin-containing monooxygenases." *Drug Metabolism and Disposition* **25**(6): 772-77.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S and Daniels S (2004). "Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents." *Molecular and Cellular Proteomics* **3**(12): 1154-69.

Roth A and Singer T (2014). "The application of 3D cell models to support drug safety assessment: Opportunities & challenges." *Advanced Drug Delivery Reviews* **69**: 179-89.

Sampaziotis F, Segeritz CP and Vallier L (2014). "Potential of human induced pluripotent stem cells in studies of liver disease." *Hepatology* **62**(1): 303-11.

Sawyer N, Worrall L, Crowe J, Waters S, Shakesheff K, Rose F and Morgan S (2008). "In situ monitoring of 3D in vitro cell aggregation using an optical imaging system." *Biotechnology and Bioengineering* **100**(1): 159-67.

Schwartz R, Fleming H, Khetani S and Bhatia S (2014). "Pluripotent stem cell-derived hepatocyte-like cells." *Biotechnology Advances* **32**(2): 504-13.

Scigelova M and Makarov A (2006). "Orbitrap mass analyzer - overview and applications in proteomics." *Practical Proteomics* **6**(Suppl 2): 16-21.

Segeritz C (2015). Investigating the pathophysiology of alpha-1-antitrypsin deficiency using human induced pluripotent stem cells Doctoral Thesis, University of Cambridge.

Selden C, Khalil M and Hodgson H (1999). "What keeps hepatocytes on the straight and narrow? Maintaining differentiated function in the liver." *Gut* **44**(4): 443-46.

Senior J (2009). "Monitoring for hepatotoxicity: what is the predictive value of liver "function" tests?" *Clinical Pharmacology and Therapeutics* **85**(3): 331-34.

Sharma NS, Wallenstein EJ, Novik E, Maguire T, Schloss R and Yarmush ML (2009). "Enrichment of hepatocyte-like cells with upregulated metabolic and differentiated function derived from embryonic stem cells using S-nitrosoacetylpenicillamine." *Tissue Engineering: Part C* **15**(2): 297-306.

Shin JW, Park SH, Kang YG and Shin JW (2012). "Potential of engineering methodologies for the application to pharmaceutical research." *Archives of Pharmacal Research* **35**(2): 299-309.

Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, Dalton S and Duncan SA (2010). "Highly efficient generation of human hepatocyte–like cells from induced pluripotent stem cells." *Hepatology* **51**(1): 297-305.

Singh RK, Gaikwad SM, Chatterjee S and Ray P (2014). Stem Cells: The Holy Grail of Regenerative Medicine. Engineering in Translational Medicine, Springer**:** 19-69.

Smith P, Krohn RI, Hermanson G, Mallia A, Gartner F, Provenzano M, Fujimoto E, Goeke N, Olson B and Klenk D (1985). "Measurement of protein using bicinchoninic acid." *Analytical Biochemistry* **150**(1): 76-85.

Song Z, Cai J, Liu Y, Zhao D, Yong J, Duo S, Song X, Guo Y, Zhao Y and Qin H (2009). "Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells." *Cell Research* **19**(11): 1233-42.

Sun LZ, Ji ZL, Chen X, Wang J and Chen YZ (2002). "ADME-AP: A database of ADME associated proteins." *Bioinformatics* **18**(12): 1699-700.

Takahashi K and Yamanaka S (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." *Cell* **126**(4): 663-76.

Takayama K, Inamura M, Kawabata K, Katayama K, Higuchi M, Tashiro K, Nonaka A, Sakurai F, Hayakawa T and Furue MK (2012). "Efficient generation of functional hepatocytes from human embryonic stem cells and induced pluripotent stem cells by HNF4α transduction." *Molecular Therapy* **20**(1): 127-37.

Temple R (2006). "Hy's law: Predicting serious hepatotoxicity." *Pharmacoepidemiology and Drug Safety* **15**(4): 241-43.

TheUniProtConsortium (2014). UniProt: a hub for protein information. Nucleic Acids Research. **43:** 204-12.

Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T and Hamon C (2003). "Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS." *Analytical Chemistry* **75**(8): 1895-904.

Touboul T, Hannan NR, Corbineau S, Martinez A, Martinet C, Branchereau S, Mainot S, Strick-Marchand H, Pedersen R and Di Santo J (2010). "Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development." *Hepatology* **51**(5): 1754-65.

Tusher VG, Tibshirani R and Chu G (2001). "Significance analysis of microarrays applied to the ionizing radiation response." *Proceedings of the National Academy of Sciences* **98**(9): 5116-21.

Tyers M and Mann M (2003). "From genomics to proteomics." *Nature* **422**: 193-97.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E and Asplund A (2015). "Tissue-based map of the human proteome." *Science* **347**(6220): 394-404.

Van Tonder JJ (2011). Development of an in vitro mechanistic toxicity screening model using cultured hepatocytes. Doctoral Thesis, University of Pretoria.

Vaudel M, Barsnes H, Berven FS, Sickmann A and Martens L (2011). "SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X! Tandem searches." *Proteomics* **11**(5): 996-99.

Vaudel M, Burkhart JM, Sickmann A, Martens L and Zahedi RP (2011). "Peptide identification quality control." *Proteomics* **11**(10): 2105-14.

Vaudel M, Burkhart JM, Zahedi RP, Oveland E, Berven FS, Sickmann A, Martens L and Barsnes H (2015). "PeptideShaker enables reanalysis of MS-derived proteomics data sets." *Nature Biotechnology* **33**(1): 22-24.

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

Vildhede A, Wisniewski JR, Noren A, Karlgren M and Artursson P (2015). "Comparative proteomic analysis of human liver tissue and isolated hepatocytes with a focus on proteins determining drug exposure." *Journal of Proteome Research* **14**(8): 3305-14.

Viner R, Scigelova M, Zeller M, Oppermann M, Moehring T and Zabrouskov V "Relative quantitation of TMT-labeled proteomes: Focus on sensitivity and precision."

Walgren JL, Mitchell MD and Thompson DC (2005). "Role of metabolism in drug-induced idiosyncratic hepatotoxicity." *Critical Reviews in Toxicology* **35**(4): 325-61.

Wang H, Chang-Wong T, Tang H-Y and Speicher DW (2010). "Comparison of extensive protein fractionation and repetitive LC-MS/MS analyses on depth of analysis for complex proteomes." *Journal of Proteome Research* **9**(2): 1032-40.

Ware BR, Berger DR and Khetani SR (2015). "Prediction of drug-induced liver injury in micropatterned co-cultures containing iPSC-derived human hepatocytes." *Toxicological Sciences* **145**(2): 252-62.

Washburn MP, Wolters D and Yates JR (2001). "Large-scale analysis of the yeast proteome by multidimensional protein identification technology." *Nature Biotechnology* **19**(3): 242-47.

Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S and Marx H (2014). "Mass-spectrometry-based draft of the human proteome." *Nature* **509**(7502): 582-87.

Wilkening S, Stahl F and Bader A (2003). "Comparison of primary human hepatocytes and hepatoma cell line HepG2 with regard to their biotransformation properties." *Drug Metabolism and Disposition* **31**(8): 1035-42.

Williams DP and Park BK (2003). "Idiosyncratic toxicity: The role of toxicophores and bioactivation." *Drug Discovery Today* **8**(22): 1044-50.

Wong C, Vosburgh E, Levine AJ, Cong L and Xu EY (2012). "Human neuroendocrine tumor cell lines as a three-dimensional model for the study of human neuroendocrine tumor therapy." *Journal of Visualized Experiments* **66**: 1-7.

Woolf TF and Jordan RA (1987). "Basic concepts in drug metabolism: Part I." *The Journal of Clinical Pharmacology* **27**(1): 15-17.

Wrzesinski K and Fey SJ (2013). "After trypsinisation, 3D spheroids of C3A hepatocytes need 18 days to re-establish similar levels of key physiological functions to those seen in the liver." *Toxicology Research* **2**(2): 123-35.

Wrzesinski K, Rogowska-Wrzesinska A, Kanlaya R, Borkowski K, Schwämmle V, Dai J, Joensen KE, Wojdyla K, Carvalho VB and Fey SJ (2014). "The cultural divide: Exponential growth in classical 2D and metabolic equilibrium in 3D environments." *PloS One* **9**(9): 1-15.

Xu JJ, Diaz D and O'Brien PJ (2004). "Applications of cytotoxicity assays and pre-lethal mechanistic assays for assessment of human hepatotoxicity potential." *Chemico-Biological Interactions* **150**(1): 115-28.

Yan W, Hwang D and Aebersold R (2008). Quantitative proteomic analysis to profile dynamic changes in the spatial distribution of cellular proteins. Organelle Proteomics, Springer**:** 389-401.

Yang X, Salminen WF and Schnackenberg LK (2012). "Current and emerging biomarkers of hepatotoxicity." *Current Biomarker Findings* **2**: 43-55.

Yusa K, Rashid ST, Strick-Marchand H, Varela I, Liu P, Paschon DE, Miranda E, Ordóñez A, Hannan NR and Rouhani FJ (2011). "Targeted gene correction of alpha-1-antitrypsin deficiency in induced pluripotent stem cells." *Nature* **478**: 391-94.

Zhang T, Jiang L, Viner R and Zabrouskov V (2015). "Improving quantitation of TMT-labeled peptides using stepped higher-energy collisional dissociation." *Thermo Fisher Scientific: Application Note 483*.

Zimmerman H (1999). "Hepatic metabolism of foreign compounds." *Hepatotoxicity* **1**: 11-40.

Ethical Approval

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.
- FWA 00002567, Approved dd 22 May 2002 and Expires 20 Oct 2016.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 22/04/2017.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

31/07/2014

Approval Certificate
New Application

Ethics Reference No.: 290/2014

Title: Proteomic assessment of potential in vitro hepatotoxicity models

Dear Ms Tracey Hurrell

The **New Application** as supported by documents specified in your cover letter for your research received on the 3/07/2014, was approved, by the Faculty of Health Sciences Research Ethics Committee on the 30/07/2014.

Please note the following about your ethics approval:
- Ethics Approval is valid for 3 years.
- Please remember to use your protocol number (**290/2014**) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:
- The ethics approval is conditional on the receipt of 6 monthly written Progress Reports, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

Dr R Sommers; MBChB; MMed (Int); MPharMed.
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

♦ Tel:012-3541330 ♦ Fax:012-3541367 Fax2Email: 0866515924 ♦ E-Mail: fhsethics@up.ac.za
♦ Web: //www.healthethics-up.co.za ♦ H W Snyman Bld (South) Level 2-34 ♦ Private Bag x 323, Arcadia, Pta, S.A., 0007